A sparsity-based variational approach for the restoration of SMOS images from L1A data

Javier Preciozzi, *Student Member, IEEE,* Andrés Almansa, *Member, IEEE,* Pablo Musé, Sylvain Durand, Ali Khazaal and Bernard Rougé

Abstract—The SMOS mission senses ocean salinity and soil moisture by measuring Earth's brightness temperature using interferometry in the L-band. These interferometry measurements known as *visibilities* constitute the SMOS L1A data product. Despite the L-band being reserved for Earth observation, the presence of illegal emitters cause radio frequency interference (RFI) that mask the energy radiated from the Earth and strongly corrupt the acquired images. Therefore, the recovery of brightness temperature from corrupted data by image restoration techniques is of major interest. In this work we propose a variational model to recover super-resolved, denoised brightness temperature maps by decomposing the images into two components: an image T that models the Earth's brightness temperature and an image O modeling the RFIs. Experiments with synthetic and real data support the suitability of the proposed approach.

Index Terms—SMOS, MIRAS, RFI, brightness temperature, non-differentiable convex optimization, total variation minimization.

I. INTRODUCTION

Observing Earth variables such as surface soil moisture (SSM) and sea surface salinity (SSS) is crucial to obtain meteorological and climate predictions [1]. The SMOS satellite [2] carries an instrument called MIRAS (Microwave Imaging Radiometer by Aperture Synthesis) [3], [4], that provides indirect measurements of the corresponding brightness temperatures of both SSM and SSS in the L-band microwave, using interferometry.

A. From brightness temperatures to visibilities: The MIRAS instrument and the forward problem

The interferometry principle used by the MIRAS instrument [3], [4] can be formalized by means of the *visibility function* [5], that relates brightness temperatures with visibilities. The antennas configuration chosen for MIRAS instrument is a Y-shaped array, where each arm is composed by a set of regularly spaced passive antennas (see Figure 1, left figure). Let (A_k, A_l) be any pair of its antennas; the visibility function

Javier Preciozzi and Pablo Musé are with the Department of Electrical Engineering, Facultad de Ingeniería, Universidad de la República, 11300 Montevideo, Uruguay (email: jprecio@dsensetechnologies.com; pmuse@fing.edu.uy).

Andrés Almansa is with CNRS MAP5, Université Paris Descartes, 75006 Paris, France (email: andres.almansa@parisdescartes.fr).

Sylvain Durand is with MAP5, CNRS & Université Paris-Descartes, 75006 Paris, France (email: sylvain.durand@parisdescartes.fr).

Ali Khazaal and Bernard Rougé are with the Centre d'Etudes Spatiales de la BIOsphère (CESBIO), 31400 Toulouse, France (email: ali.khazaal@cesbio.cnes.fr; bernard.rouge@cesbio.cnes.fr).

$V(\mathbf{u}_{kl})$ is defined as the complex cross-correlation between the received signals at A_k and A_l:

$$V(\mathbf{u}_{kl}) = \frac{1}{\sqrt{\Omega_k \Omega_l}}$$
$$\iint_{||\boldsymbol{\xi}|| \leq 1} U_k(\boldsymbol{\xi}) U_l^*(\boldsymbol{\xi})(T_b(\boldsymbol{\xi}) - T_r)\tilde{r}_{kl}(t)\frac{e^{-i2\pi\mathbf{u}_{kl}^T\boldsymbol{\xi}}d\boldsymbol{\xi}}{\sqrt{1 - ||\boldsymbol{\xi}||^2}}. \quad (1)$$

Here U_k and U_l are the corresponding normalized voltage patterns of the antennas; Ω_k, Ω_l their corresponding solid angles; $\mathbf{u}_{k,l} = (u_1^{kl}, u_2^{kl})$ is the spacing between the two antennas (A_k, A_l) in wavelength. The Cartesian coordinates $\boldsymbol{\xi} = (\xi_1, \xi_2)$ are the spatial domain coordinates, which are restricted to the unit circle. T_r is the physical temperature of the receivers (assumed the same for all receivers); $\tilde{r_{kl}}$ is the Fringe-Wash function, a function of the spatial delay $t = \frac{\mathbf{u_{kl}}^T \boldsymbol{\xi}}{f_0}$, where $f_0 = \frac{c}{\lambda_0}$ is the central frequency of observation. Note that the brightness temperature T_b is a 2D function defined on the unit circle $\{\boldsymbol{\xi} : ||\boldsymbol{\xi}|| \leq 1\}$ [5], [6].

Since the brightness temperature support is the unit circle, it is well known that the best regular sampling grid is a hexagonal one [7], leading to the largest alias-free Field Of View for a given spacing between the antennas [8]. The antennas configuration chosen for the MIRAS instrument is a Y-shaped array [8], shown in Figure 1 along with its corresponding hexagonally sampled, star-shaped domain H. This star-shaped domain H, known in aperture synthesis as the *experimental frequency coverage* [9, Ch. 5], is contained within the hexagon, hence requiring extrapolation to recover the missing parts (see Figure 2). Because of other practical
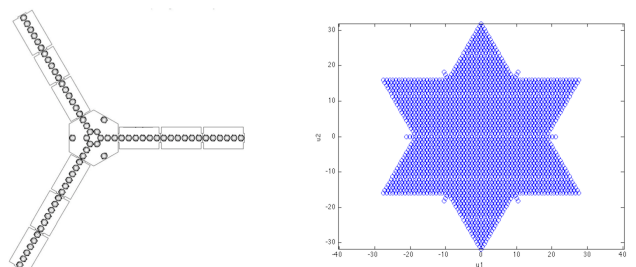

Fig. 1. Diagram of the MIRAS instrument onboard of SMOS satellite (from [10]) and the star-shaped, hexagonally sampled visibilities domain derived from it. The number of antennas in the miras instrument is 69: 22 antennas type Light-Weight Cost-Effective Front-End (LICEF) in each arm, plus three antennas of type Noise Level Injection Radiometers (NIR) located in-between each arm in the center of the instrument.

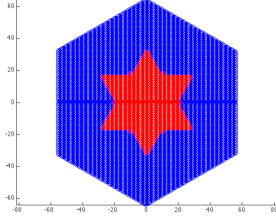issues, the final separation between antennas in the MIRAS

Fig. 2. The absence of information beyond the *experimental frequency coverage H* (in red) makes the recovery of the brightness temperature map an ill-posed problem.

instrument is $d := \|\mathbf{u}_{kl}\| = 0.875$ wavelengths, for any $k \neq l$ within each arm. This value is larger than the critical sampling distance on a hexagonal grid (which is $d = 1/\sqrt{3}$ wavelengths), and then, some aliasing is introduced on the spatial domain. Figure 3 shows the spatial domain and its six replicas, where the Alias-Free Field of View is the small central zone in the hexagon delimited by the blue dashed lines. Because of the tilt of the satellite acquisition system, most of the aliasing comes from the intersection between the Earth disk with the known cold sky (whose values are known) and therefore it is customary to consider the Field of View delimited by the red plain lines: the *Extended Alias-Free Field of View* (E-AF-FOV) [8].

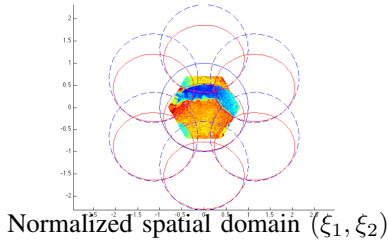Normalized spatial domain (ξ_1, ξ_2)

Fig. 3. Alias Free - Field of View (AF-FOV) – zone delimited by the unity circle contours in blue lines – and the corresponding Extended Alias-Free Field of View (E-AF-FOV) in red lines. See text for details.

Remark: As one can readily see from Equation (1), brightness temperatures T are originally defined on the unit circle. However things are quite different in the discrete setting $\mathbf{G}T = V$. In this setting, as already mentioned, the visibilities $V(u)$ can only be measured for discrete values of the frequencies $u = u_{kl}$ which are given by the discrete spacings between available antennas. Hence, in the discrete setting the matrix $\mathbf{G}$ is a discretization of the linear operator $V = G(T)$ defined by Equation (1).
Each row of the matrix $\mathbf{G}$ is associated to Equation (1) for one particular value of the frequency u_{kl}.
Each column of the matrix $\mathbf{G}$ is associated to a particular value of the spatial variable ξ.
The row discretization is dictated by the arrangement of the antennas that allow to measure the visibilities only for discrete values of u_{kl}.
The column discretization is dictated by the sampling scheme used for T in the spatial domain, and can in principle be chosen at will. In our work we made the standard choice of

defining T on a regular (hexagonal) spatial grid that is based on the dual grid of the hexagonal grid where the u_{kl}s are defined. This way we can use the standard FFT algorithm to compute discrete Fourier transforms between the two grids.

(Inevitable) row discretization leads to spatial aliasing: The u_{kl}s however are defined on a hexagonal sampling grid, supported on a star-shaped domain (the so called experimental frequency coverage) resulting from the Y-shaped antennas array.
As explained above the spacing between antennas is not fine enough to ensure a larger-that-critical frequency sampling. The result is spatial aliasing that we do not pretend to be able to remove by our method. Instead we only keep as reliable information the so-called Extended Alias Free Field of View (EAF-FOV) area which is not affected by spatial aliasing artifacts (see Figure 3). However computations are performed on a full hexagon which is larger than the EAF-FOV. The full visible circular spatial domain is larger than this hexagon (see Figure 3), but extending the computations to the full circle would be useless because it would contain periodic copies of the hexagonal tile. In fact, since acquisition happens at a discrete grid in the frequency domain, the recovered image is periodic in the spatial domain.

Finer column discretization and TV minimization improve resolution of T. The sampling rate in the spatial domain can be chosen at will. A finer sampling rate means more columns in matrix $\mathbf{G}_{kl}$, and a more accurate Riemann sum approximation of the integral in Equation (1). The spatial sampling resolution we used corresponds to the hexagonal reciprocal cell shown in Figure 1. The standard L1B product only produces Fourier coefficients restricted to the red star shaped domain in Figure 1. The output of our Algorithm produces a brightness temperature map whose Fourier transform is larger than the star-shaped domain. The additional Fourier coefficients are inferred from a reasonable piece-wise regular prior for brightness temperatures. This prior (implemented by total variation minimization under sensing constraints) allows to better resolve the sharp boundaries between land and sea temperatures for instance.

B. Restoring brightness temperatures from visibilities: An ill-posed inverse problem

Going back to Equation (1), after subtracting the contribution of T_r using the flat target response, the samples of T in the hexagonal grid can be obtained from the visibility samples by solving the linear system $\mathbf{G}T = V$, where matrix $\mathbf{G}$ represents the discrete linear operator given by (1). Of course, the inversion of this problem is ill-posed since $\mathbf{G}$ is not invertible. The ill-posedness of the problem can also be interpreted as a lack of information beyond H, as illustrated in Figure 2. Hence, additional constraints must be added to the model. In [6], the authors propose to solve it as a constrained least square minimization problem, imposing that T has no frequency components outside H. This problem can be formulated as an unconstrained minimisation:

$$\hat{T}_{sol} = \min_{\hat{T}} \|V - \mathbf{G}\mathbf{F}^*\mathbf{Z_H}\hat{T}\|_2^2 \qquad (2)$$

where $\mathbf{F}^*$ denotes the matrix corresponding to the hexagonal Inverse Fourier Transform, $\mathbf{Z_H}$ the zero padding operator beyond H and $\hat{T}$ the Fourier coefficients of T for frequencies in H. This minimisation problem has a direct solution: $\hat{T} = \mathbf{J}^+ V$ where $\mathbf{J} = \mathbf{GF}^*\mathbf{Z_H}$ and $\mathbf{J}^+ = (\mathbf{J}^*\mathbf{J})^{-1}\mathbf{J}^*$ is the pseudo-inverse of $\mathbf{J}$. Finally, T can be obtained from $\hat{T}$ very easily applying an inverse Fourier transform. This approach is the nominal image reconstruction process of the SMOS pipeline and corresponds exactly to the L1B product. It is well known that a zero padding on the Fourier domain may lead to Gibbs effects on the restored image. This can be partially alleviated (as proposed by [6]) by the use of a Blackman window $\mathbf{B}$: $T = \mathbf{F}^*\mathbf{BZ_\Omega}\hat{T}$.

This approach (that we will refer in what follows as the *zero padding* approach), has many advantages: it has a clear physical interpretation (values outside H are set to zero), it is very fast (since only a matrix multiplication is needed to obtain $\hat{T}$), and in general gives good results. Nevertheless, this is not the case when the visibilities are corrupted by radio frequency interferences (RFI). Because these RFIs have frequencies beyond H and their power is far stronger than Earth radiation, very strong Gibbs effects can be seen on the final *brightness temperature* images (see for instance the bottom left image in Figure 10, corresponding to western Europe, obtained by the nominal reconstruction process [6]).

C. Previous work

Several methods for post-processing SMOS images reconstructed using the nominal method have been proposed in the literature. However, most of them have been directed to both mitigation and flagging of RFIs from L1B data. In [11], all pixels whose brightness temperature exceeds threshold of 350 K (the maximum, blackbody radiation at 1.4 GHz) are flagged as possible RFIs.Then, several snapshots are analysed and statistical information of the RFI is computed. If the number of RFI on the given geolocation is larger than a certain number (40 in the cited article), the point is marked as a RFI and the most probable brightness temperature is estimated. With this estimation, a process similar to the Sun correction [12] is performed: the visibility of the RFI is computed using matrix $\mathbf{G}$ and finally this visibility is subtracted from the original one. In [13], the authors generate synthetic brightness temperature images that are as close as possible to the natural ones. Then, the corresponding visibilities are computed and subtracted from the original ones. If RFIs are present in the scene, they will be clearly noted on the brightness temperature obtained from this subtraction. The same thresholding principle as in [11] is applied, obtaining possible RFIs locations, but a gradient threshold is applied as well. After that, a local optimisation process taking into account the Point Spread Function of the instrument is applied on a surrounding region, to better localize each RFI. Finally, in [14] the authors focus their attention on the accurate geolocation of RFI emitters, since its main objective is to take political actions on them. Here again, a thresholding on the brightness temperatures

similar to [11] is applied and a clustering process is performed with the obtained potential RFI candidates. On each cluster, an interpolation on a finer grid is performed, thus obtaining a better geolocation of the RFI emitter.

However, to our knowledge and as pointed out in a recent work [15], since the nominal method proposed by Anterrieu et al. [16] in 2002, not much effort has been done in providing fully restored images in the sense of solving the inverse problem given by (1). González-Gambau et al. [15] propose a new restoration method based on the underlying hypothesis that the geophysical signal of interest varies slowly, and in that case they manage to reduce the impact of RFIs in the reconstruction process, in particular the highly oscillatory patterns they introduce. This method is particularly useful to improve the salinity estimation in the ocean, however the authors have reported that some RFI sources present a widening effect with respect to the result obtained by the nominal method. Moreover, since on land applications the geophysical signal varies faster, the approach presents low performance.

D. Proposed approach and contributions

In this work we propose a novel approach to solve the inverse problem presented on (1) using a variational formulation that explicitly models the formation of visibilities as a superposition of the Earth's natural brightness temperatures and the RFIs. As it will be demonstrated later, the proposed method automatically removes signal effects generated from RFIs, while at the same time extrapolates the image spectrum in order to minimize Gibbs effects, with no *a priori* knowledge of the RFI's locations. A preliminary, short version of this work was presented in IGARSS 2014 [17].

The article is organized as follows. In Section II we present the model on which our restoration approach is based, and we explain how this model can be solved using state of the art optimization techniques. In Section III we describe in detail numerical issues to make the problem numerically tractable. The definitive implementation is presented in Section IV, where we discuss major issues like parameter choices and numerical optimization. In Section V we present experiments with synthetic data that validate the proposed approach and we also apply our method to real L1A SMOS data and we compare them with results obtained using the *zero padding* and the *Blackman* approach. For the sake of brevity, a large set of experiments is presented in the complementary material. Conclusions and future work are presented in Section VI.

II. MODELIZATION AND VARIATIONAL FORMULATION: A FIRST APPROXIMATION TO THE PROBLEM

Visibilities and *brightness temperatures* are related by the linear operator given by (1). In its discrete form, using matrix notation[1], this is $\mathbf{G}T = V$. As before, the goal is to obtain the original temperatures image T from the given visibilities V,

---

[1]For the sake of simplicity, we use the same notation to refer to an image and its vectorized form. Disambiguation follows easily from the context.

knowing that, as noted before, $\mathbf{G}$ is not invertible. We propose to model the observed brightness temperature image $\tilde{T}$ as

$$\tilde{T} = T + O, \tag{3}$$

where T is the non-polluted Earth's *brightness temperature* image and O is the outliers' image generated by the RFIs that are assumed to be sparse pointwise sources (delta functions). Because of the physical acquisition system, the visibilities are corrupted with noise, that under realistic assumptions can be considered to be a Gaussian white noise with 0 mean and standard deviation $\sigma_{\Delta V_L}$ [18]. The final equation is then:

$$\mathbf{G}\tilde{T} + n = \mathbf{G}(T + O) + n = V \tag{4}$$

where V is the vector of measured visibilities (corrupted by the additive Gaussian noise n described before), after removing the unwanted contributions from the sky, sun and backlobes [19].

A. Variational formulation

We propose to recover T by solving the following constrained minimization problem:

$$\begin{aligned} (T_{sol}, O_{sol}) = \min_{T,O} \{ \text{TV}(T) + \mu S(O) \} \\ \text{such that } \|\mathbf{G}(T + O) - V\|_2^2 \le |H|\sigma_{\Delta V_L}^2. \end{aligned} \tag{5}$$

In this formulation, $TV(\cdot)$ denotes the total variation seminorm (Appendix A) and $S(\cdot)$ is a norm that promotes sparsity of the outliers image O_{sol}, for instance the ℓ_1 norm or the ℓ_0 counting measure (Appendix A). Parameter μ controls the trade-off between both terms; its choice can be formally derived from geometric considerations on the outliers, and will be discussed in detail in Section IV-A. The data fit term is derived directly from the white, Gaussian distributed data noise model $n = \mathbf{G}(T + O) - V$. The bound $|H|\sigma_{\Delta V_L}^2$ represents the area of H times the visibilities variance $\sigma_{\Delta V_L}^2$.

It is well known that the minimization of the ℓ_2 norm corresponds to the *Maximum A Posteriori* estimator of white Gaussian noise. In the SMOS instrument, the visibilities are always perturbed by a random radiometric noise, which is assumed to be uncorrelated (on the visibilities) and distributed according a Gaussian model with standard deviation [18]:

$$\sigma_{\Delta V_L} = \frac{T_A + T_R}{\sqrt{2B\tau}}, \tag{6}$$

where T_A is the antenna temperature, T_R is the receiver temperature, B is the bandwidth and τ is the integration time. The bandwidth for MIRAS instrument is $B = 19$ MHz and the integration time is $\tau = 0.663$ seconds[2]. The other two variables (T_A and T_R) depend on the observed scene and operational conditions. The value of those temperatures was reported to be $T_A = 294$ K and $T_R = 200$ K, leading to a noise standard deviation $\sigma_{\Delta V_L} = 0.098$ K.

Problem (5) can be reformulated as an unconstrained one:

$$(T_{sol}, O_{sol}) = \min_{T,O} \|\mathbf{G}(T + O) - V\|_2^2 + \lambda(\text{TV}(T) + \mu S(O)); \tag{7}$$

where the Lagrange multiplier λ must be chosen to ensure equivalence between both problems. Parameter λ is therefore the counterpart of the noise constraint in (5), and it is well known that there exist a unique Lagrange multiplier λ, uniquely determined by $|H|\sigma_{\Delta V_L}^2$, ensuring that both the constrained and unconstrained problems admit the same solution. A classical way to set λ programmatically is to use Uzawa's algorithm [20], described on Algorithm 2. Basically, the procedure consists in starting with a fixed λ^0, then solving the original problem with this λ^0 and adjusting λ^k iteratively, depending on how near or how far the solution is to the original constraint. The procedure continues with ensured convergence until the data fit term is close enough to the expected noise $|H|\sigma_{\Delta V_L}^2$.

Remark: The total variation is used to super-resolve T beyond the *experimental frequency coverage* H while avoiding Gibbs oscillations: it is a complete substitution for the *zero padding* approach, where instead of filling the coefficients outside H with zero, they are recovered by the TV minimization [21]. On the other hand, the sparsity operator is chosen to model the RFIs image O since this image is zero almost everywhere because RFIs are very sparsely located. It is well known that the ℓ_1 or the ℓ_0 norms promote sparsity [22]: consequently, image O_{sol} is expected to recover a sparse image composed only by the RFIs.

When $S(\cdot)$ is the ℓ_1 norm, problem (7) is convex, and despite not being strictly convex, it admits a unique local (and therefore global) minimizer ([23], [24]). However, a careful treatment has to be taken since the term multiplied by λ is non-differentiable. Fortunately, there exist optimization methods to solve this kind of problems. One of them is the *Forward-Backward* (FB) splitting algorithm [25], while other splitting approaches such as the Split-Bregman method could also be used[3]:

Proposition 1. *Let $E(x) = E_1(x) + E_2(x)$, where E_1 and E_2 are convex functions such that E_1 is differentiable with Lipschitz gradient, and E_2 is a* simple *function, in the sense that its associated proximal operator*

$$\text{prox}_{\gamma E_2}(x) = \arg\inf_y \ E_2(y) + \frac{1}{2\gamma}\|x - y\|^2 \tag{8}$$

admits a closed form or a simple algorithm to compute it. If these hypotheses hold, the following generic algorithm can be derived: for each $k \in \mathbb{N}$, the k-th iteration starting from seed $x^0 = (T^0, O^0)$ is given by

$$\begin{cases} x^{k+1/2} & = & x^k - \gamma \nabla E_1(x^k) \\ x^{k+1} & = & \text{prox}_{\gamma E_2}(x^{k+1/2}). \end{cases}$$

In order to ensure convergence to the minimizer, γ must be smaller than $2/L$, where L is the Lipschitz constant of ∇E_1.

---

[2]Integration time is in fact 1.2 seconds, but it is adjusted by a factor of 1.81 to take into account correlation and sampling effects.

[3]We decided to use FB since, while the parameters in FB follow naturally from the implementation, those involved in the SB algorithm are related to convergence rates and are therefore harder to set.

In our case, we have $E_1(T, O) = \|\mathbf{G}(T + O) - V\|_2^2$ and $E_2(T, O) = \lambda(\mathrm{TV}(T) + \mu S(O))$ that yields

$$\nabla E_1(T, O) = \left(\begin{array}{c} \mathbf{G}^*\mathbf{G}(T + O) - V \\ \mathbf{G}^*\mathbf{G}(T + O) - V \end{array} \right),$$

$$\mathrm{prox}_{\gamma E_2}(T, O) = \left(\begin{array}{c} \mathrm{prox}_{\gamma\lambda\,\mathrm{TV}}(T) \\ \mathrm{prox}_{\gamma\lambda\mu\|\cdot\|_1}(O) \end{array} \right).$$

It is straightforward to show from (8) that the $\mathrm{prox}_{\gamma\lambda\mu\|\cdot\|_1}(O)$ corresponds to the *soft thresholding* operator (see Appendix A):

$$s_\Delta(t) = \left\{ \begin{array}{ll} \mathrm{sign}(t)(|t| - \Delta) & \text{if } |t| \geq \Delta, \\ 0 & \text{if } |t| < \Delta. \end{array} \right.$$

Although $\mathrm{prox}_{\gamma\lambda\,\mathrm{TV}}(T)$ does not have an explicit form, there exist several algorithms that solve this proximal operator ([23], [24]). In Appendix A we explain the one used in this work.

The algorithm described so far converges to the unique global minimizer that corresponds to the solution of problem (5) with sparsity operator $S(O) = \|O\|_1$. We use this solution as an initialisation for the second step, where the sparsity operator is chosen to be $S(O) = \|O\|_0$, which is non-convex. For this problem, the same FB algorithm can be considered and is guaranteed to converge to a local minimizer [26]. Now, instead of the soft thresholding, the proximal operator for $S(O) = \|O\|_0$ becomes the hard thresholding $h_{\sqrt{2\gamma\lambda\mu}}(t) = t\,\mathbb{1}_{\{|t| \geq \sqrt{2\gamma\lambda\mu}\}}$ (see Appendix A). The reason to run the FB algorithm with $\|O\|_0$ is to exploit even more the sparsity characteristics of the problem, since it is known that norm $\|O\|_0$ leads to a sparser solution than $\|O\|_1$ [22].

Note that in the case of ℓ_0 norm, the convergence of the Uzawa's algorithm is not guaranteed. Thus, in this second step, we do not apply Uzawa's algorithm: we consider that the λ value obtained for ℓ_1 is correct also for the ℓ_0 norm. In fact, because the solution obtained with ℓ_1 is a very good approximation to the sparse representation, a few iterations of the FB with ℓ_0 is enough to converge to the local minimum.

III. NUMERICAL ISSUES

A. Matrix $\mathbf{G}^*\mathbf{G}$

The FB algorithm requieres a multiplication by matrix $\mathbf{G}^*\mathbf{G}$ at each iteration. Matrix $\mathbf{G}$ is a 4695×16384 matrix: 4695 is the number of visibilities and 16834 is the number of samples of the *brightness temperature*. It follows that matrix $\mathbf{G}^*\mathbf{G}$ is 16384×16384. Because it is a dense matrix, a great number of operations are performed at each iteration, which is computationally not desirable. However, a change of basis to the Fourier domain yields:

$$\nabla E_1(T, O) = \mathbf{F}^*((\mathbf{G}\mathbf{F}^*)^*\mathbf{G}\mathbf{F}^*\mathbf{F}(T + O) - (\mathbf{G}\mathbf{F}^*)^*V), \quad (9)$$

where $\mathbf{F}$ is the matrix representation of the Fourier transform (and $\mathbf{F}^*$ its conjugate) . The advantage of this representation is that matrix $((\mathbf{G}\mathbf{F}^*)^*\mathbf{G}\mathbf{F}^*$ reveals a highly sparse structure: to keep the Frobenius norm of $((\mathbf{G}\mathbf{F}^*)^*\mathbf{G}\mathbf{F}^*$ at 99.99% of its value, we only need to keep 0.0008 of its entries. From a computational point of view, solving Equation 9 can be done

even more efficiently using a standard Fast Fourier Transform implementation:

$$\nabla E_1(T, O) = \mathcal{F}^*((\mathbf{G}\mathbf{F}^*)^*\mathbf{G}\mathbf{F}^*\mathcal{F}(T + O) - (\mathbf{G}\mathbf{F}^*)^*V), \quad (10)$$

where $\mathcal{F}$ denotes the Cooley and Tuckey standard Fast Fourier Transform, that we use instead of matrix multiplication by the matrix $\mathbf{F}$.

The property of being sparse on the Fourier domain can be easily understood from the image formation point of view: each of the visibilities values corresponds to the cross-correlation of two signals received at two different antennas. Clearly, if the antennas are close to each other this correlation is high, but the farther they get it tends to be negligible.

The use of the standard (rectangular) FFT on a hexagonal lattice can be made possible by re-projecting the samples in a rectangular grid (see [8] or [16, Section III]). Consequently, we can modify the FB algorithm described above to perform the minimization in the Fourier domain. This implementation is summarized in Algorithm 1 (where all the variables have the same meaning as before).

Proposed method with Fordward-Backward implementation

input : A vector of visibilities V
output: A brightness temperature image T, and an RFIs image O
initialization:
Step 0:
- Set $T^0 = 0$ and $O^0 = 0$
- Choose $\gamma < 2/L$, where L is the Lipschitz constant of $\nabla(E_1)$

Step k ($k \geq 0$):

$$
\begin{aligned}
T^{k+\frac{1}{2}} &= T^k - \gamma\mathcal{F}^*(\mathbf{F}\mathbf{G}^*\mathbf{G}\mathbf{F}^*\mathcal{F}(T^k + O^k) - \mathbf{F}\mathbf{G}^*)V \\
O^{k+\frac{1}{2}} &= O^k - \gamma\mathcal{F}^*(\mathbf{F}\mathbf{G}^*\mathbf{G}\mathbf{F}^*\mathcal{F}(T^k + O^k) - \mathbf{F}\mathbf{G}^*)V \\
T^{k+1} &= \mathrm{prox}_{\gamma\lambda\,\mathrm{TV}}(T^{k+1/2}) \\
O^{k+1} &= s_{\gamma\lambda\mu}(O^{k+1/2}).
\end{aligned}
$$

Algorithm 1: Proposed method with Fordward-Backward implementation in the Fourier domain.

B. Spectral TV

In order to reduce the "staircasing" effect inherent to many TV minimization methods([27], [28]), we use an approach inspired on the *Spectral TV* method. This idea was first introduced in [21] and further developed by Moisan [29], from which our implementation is derived. In this approach, staircaising reduction is achieved by: (i) Computing image derivatives not by finite differences but analytically on the Fourier series expansion; (ii) Approximating the continuous TV as a Riemann summation over a grid at least two times finer than the critical sampling rate. In our case, instead of doubling the sampling rate, we choose to extend the spectral domain of T to an intermediate cell $\mathcal{H}$, in-between the *experimental frequency coverage* H, and the cell $\mathcal{C}$ corresponding to

the (largely overcritical) spatial sampling rate of T (Figure 4). There is another reason to use the *Spectral TV*. SMOS data
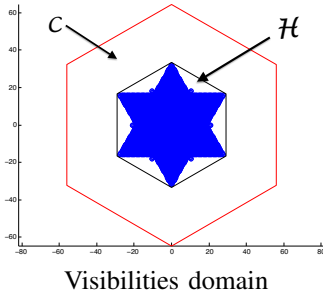

Fig. 4. Intermediate, hexagonally shaped cell $\mathcal{H}$ used to reduce the staircaising effect inherent to discrete TV minimisation methods.

is captured on a hexagonal grid, and thus the variables T and O are modeled on a hexagonal grid as well. Most TV algorithms are based on the computation of derivatives using discrete differences, that are not straightforward to adapt to the hexagonal grid. This problem is avoided if we use the *Spectral TV*, because it computes the image derivatives analytically in the Fourier expansion. In Appendix A we introduce the modifications done to Chambolle's algorithm in order to implement the *Spectral TV*. An alternative implementation based on the work presented in [30], was recently proposed in [31] and in [32, Ch. 3]. In what follows, we denote the *Spectral TV* of an image T based on the domain $\mathcal{H}$ as $\text{TV}_{\mathcal{H}}(T)$.

IV. FINAL IMPLEMENTATION

A. On the choice of μ, a resolution parameter

Parameter μ is actually the only user parameter of our method. It controls the trade-off between $\text{TV}_{\mathcal{H}}(T)$ and $S(O)$. Depending on its value, a visual structure on the scene will be considered as an outlier (and therefore, assigned to the outliers image O) or as part of the scene (the Earth *brightness temperature* image T). Its value can be obtained as follows. Consider a cylinder c of radius r and height h. The involved norms or semi-norms are

$$TV(c) = 2\pi r h, \quad \|c\|_0 = \pi r^2 \mathbb{1}_{[h>0]}, \quad \|c\|_1 = \pi r^2 h.$$

When the sparsity operator is the ℓ_1 norm, c is considered as an outlier if $TV(c) \geq \mu\|c\|_1$, leading to $\mu \leq \frac{2}{r}$. In the experiments, we have selected a value $\mu \simeq \frac{2}{10} = 0.2$, which amounts to consider that the radii of the outliers are at most 10 pixels wide. However, Earth structures spanning less than 10 pixels, can still be recovered by choosing a larger value for μ. It is worth noting that different values for μ can be used in order to have a fixed support in spatial resolution at no extra computational cost, by using what area represents a pixel at each image location.

In the case of the ℓ_0 norm, a cylinder is considered as an outlier as soon as $TV(c) \geq \mu\|c\|_0$, i.e. $\mu \leq \frac{2h}{r}$. Hence, μ can then be interpreted as the minimal h/r outlier ratio. Here, contrarily to the ℓ_1 case, the height (here expressed in Kelvin units) plays an important role in distinguishing an outlier from the data, and larger values for μ may be chosen ($\mu \simeq 100\frac{2}{10} = 20$).

In this sense, μ can be viewed as a scale parameter, and as such it is unavoidable. Any other restoration method involves the choice of similar parameters, on one way or the other. For instance, besides the intrinsic unavoidable instrument bandwith, the cut-off frequency in the zero padding method imposes a resolution limit (and produces a strong Gibbs effect as well). The same holds for the Blackman apodization method, where the resolution limit is determined by the cut-off frequency of the Blackman window. The advantage of our method is that parameter μ, as described above, has a clear geometric meaning and its choice is very intuitive. The effect of these parameters will become clear once we will have discussed Figure 5 in Section V.

Uzawa

input : L1A SMOS Data - Visibilities: $\tilde{V}$
input : Uzawa tolerance step: u_{tol}
output: A temperature image T, an outliers image O
initialization:
Step 0: Set $T^0 = 0$ and $O^0 = 0$. Set λ^0 with any suitable value (for instance, 1).

while $|u_s| > u_{tol}$ **do**
$\quad (T^{k+1}, O^{k+1}) =$
$\quad \min_{T,O} \|\mathbf{G}(T + O) - \tilde{V}\|_2^2 + \lambda^k (\text{TV}(T) + \mu S(O))$;
$\quad u_s = \|\mathbf{G}(T^k + O^k) - \tilde{V}\|_2^2 - \sigma_{\Delta V_L}^2$;
$\quad \lambda^{k+1} = \max\{\lambda^k + \rho u_s, 0\}$;
end

Algorithm 2: Uzawa general algorithm applied to find the correct value λ that is consistent with $|H|\sigma_{\Delta V_L}^2$ in the corresponding constrained minimization problem. Parameter ρ is set once for all to control the convergence rate. See [20] for further details.

B. Numerical optimization

Several methods have been proposed to accelerate the FB convergence rate. Two related approaches are FISTA [33] and monotone-FISTA [34]. The main difference between both methods and the FB algorithm is a clever combination of the two previous iterations, that increases the convergence rate. FISTA algorithm does not ensure monotone convergence (neither do FB) , which is not a desirable property for an optimization algorithm. In [34], the authors introduce a monotone version of this algorithm known as MFISTA, that does not change the convergence rate of the original FISTA algorithm, but guarantees monotonicity. The final and complete algorithm is presented in Algorithm 3 but can be summarized as follows. It consists of two loops: an external loop that implements the *Uzawa* algorithm, solving the problem of finding the correct Lagrange multiplier λ. The internal loop implements de MFISTA algorithm, which is basically the FB algorithm with additional auxiliary variables to accelerate the convergence rate.

V. EXPERIMENTAL RESULTS

In this section we present results obtained on both simulated and real data. In both cases, we compare our approach to the

---

FINAL PROPOSED ALGORITHM

input : L1A SMOS Data - Visibilities: $\tilde{V}$
input : Uzawa tolerance step: τ_u
input : MFista tolerance step: τ_m
output: A brightness temperatures image T, an RFI
image O

initialisation: Set $T^0 = 0$ and $O^0 = 0$;
Initialize Uzawa: Set $\lambda^0 = 1$ and $\rho \in (0..1]$;
while $|e_u| > \tau_u$ **do**

   Initialize MFISTA: Set $t^0 = 1$ and $e_m > \tau_m$;

   while $e_m > \tau_m$ **do**

      $(T_z^k, O_z^k) = \text{prox}_{\frac{\gamma}{L} E_2}(T^k, O^k) - \frac{\gamma}{L}\nabla E_1(T^k, O^k))$

      $t^{k+1} = \frac{1+\sqrt{1+4(t^k)^2}}{2}$

      $(T_x^k, O_x^k) =$
      $\arg\min_{(t,o)=\{(T_z^k,O_z^k),(T_x^{k-1},O_x^{k-1})\}} E(t,o)$
      $(T^{k+1}, O^{k+1}) =$
      $(T_x^k + \frac{t^k-1}{t^{k+1}}(T_z^k - T_x^{k-1}), O_x^k + \frac{t^k-1}{t^{k+1}}(O_z^k - O_x^{k-1}))$
      ;

      $e_m = |E(T^{k+1}, O^{k+1}) - E(T^k, O^k)|$

   end

   $e_u = \|\mathbf{G}(T^k + O^k) - \tilde{V}\|_2^2 - \sigma_{\Delta V_L}^2$;
   $\lambda^{k+1} = \max\{\lambda^k + \rho e_u, 0\}$;

end

Algorithm 3: Final proposed algorithm

---

nominal reconstruction process: the zero-padding approach [6] and the Blackman apodization approach [2].

A. Simulated Data

In order to have an overall idea of the proposed method's performance, we analyze it on different contexts:

- No RFIs are present in the image: only instrumental noise;
- RFIs with intensities in the same range as the Earth's brightness temperature plus instrumental noise;
- RFIs with intensities on a larger range $((0,35000] \text{ K})$[4] plus instrumental noise;
- RFIs located outside the image grid (with sub-pixel accuracy in non-integer positions) plus instrumental noise;
- Only the synthetic Earth's brightness temperature is considered (no instrumental noise or RFIs). This scenario is considered to evaluate the extrapolation power of the proposed approach.

The process to generate the simulated images is the following:

1) We generate a *brightntess temperatures* image T_{gt}, based on [35], where gt stands for *ground truth*;
2) We obtain the visibilities associated to image T_{gt}: $V_{gt} = \mathbf{G}T_{gt}$;
3) We generate a visibility noise random vector n, with independent identically distributed Gaussian entries with zero mean and covariance matrix $\sigma_{\Delta V_L}^2 \mathbf{Id}$;

---

[4]This range corresponds to the range we have observed on real data

4) We add the noise visibilities to the ground truth visibilities to obtain an outlier-free noisy visibility vector: $V = V_{gt} + n$;
5) We generate a set of RFIs' images, one for each of the cases described above (no RFIs, RFIs in the image range, RFIs between $(0,35000]$ K and RFIs outside the grid positions). If we note T_δ any of these synthetic RFIs' images, its corresponding induced visibilities is: $V_\delta = \mathbf{G}T_\delta$;
6) Finally, we obtain the simulated visibilities adding both terms: $V_f = V + V_\delta = V_{gt} + V_\delta + n$, which are in full agreement with the image formation model.

The RFIs positions and intensity values are generated randomly, and the number of outliers varies from 2 to 10[5].

Error measurement: In order to obtain a quantitative analysis the results, we need to define an error measure. Following [6], we consider the computation of the RMS error between the original image T and the recovered image T_r, over the entire Extended Alias-Free Field of View (E-AF-FOV). Because of the band-limited property of the MIRAS instrument, even in an ideal case with no instrumental noise, we could not restore the original image T: only a smoothed version of T can be recovered. In [6], results are compared against a smoothed version of T, obtained by an inverse fourier Transformation considering only the coefficients inside the the *experimental frequency coverage*, denoted by T_w. Because we want to analyze the extrapolation properties of the presented method, we report the RMS error using both T_w and T: the RMS error computed against T_w describes the recovering capabilities of our method without considering the spectral extrapolation contribution, while the RMS error computed against T provides also this information.

Figure 5 shows the result of our method when applied to an image with no RFIs, but with instrumental noise. We recall that no modification to the functional is needed: we still consider the sparsity operator on a the outliers image O, whose values at the end of the minimization are all zero. Although the RMSE against the band-limited image T_w is more or less the same for all methods, it is clear that the proposed method outperforms the others when looking at the RMSE error computed against the original image T. This is a clear improvement, due mainly to the spectral extrapolation, that can also be seen on the edges of the difference images. This experiment confirms that the proposed method is generic, in the sense that it can be applied to images without RFIs. band-limited

From this experiment it is also clear the decrease on spatial resolution due to the lowpass characteristics of MIRAS instrument: the band-limited image T_w is contaminated with Gibbs oscillations and Menorca and Ibiza-Formentera islands are severely smoothed. The three rows of Figure 5 show the results obtained using zero padding, zero padding with Blackman apodization, and the proposed method. Note that the latter yields a sharper image, with lowest RMSE. Note also that although the proposed method fails to completely

---

[5]This is not a limitation: it is based on the observation that rarely a real image has more than 10 outliers.

recover all the structures in the scene (for instance, Menorca is still severely smoothed as a result that its intensity level is close to the ringing appearing in T_w), the result is still better than the ones obtained with the zero padding or the nominal method.

Figures 6 and 7 show the results obtained when RFIs are present in the scene. In the first case, the RFIs intensities fall within the Earth's brightness temperature range; in the second case, the RFIs intensities range from 0 to 35000 K. Here again, the proposed method is consistently better than the zero padding approach and even in the presence of very large outliers (35000 K) the maximum error is 200 times smaller, leaving an error of near 150 K, which is in the order of the ground truth image values. In these examples the loss of small structures becomes more evident: all Baleares islands were removed.

In Figure 8 we present the results obtained when the outliers are not located on the sampling grid points. It can be observed that the performance of the proposed method continues to be consistent with the case where the outliers were located on the grid.

Table I summarizes the quantitative measures corresponding to the results obtained with the proposed method in all the simulated scenarios. It is clear that the method behaves consistently in all of them. We can conclude that the method is well adapted for several contexts that may occur on real data, and that no context-dependent strategies are needed in order to apply it. In other words, the method and its parameters can be set once for all, independently on the observed data.

We end this set of experiments on synthetic data with an analysis of the spectral extrapolation capability of the proposed method. In this test, we only analyse the results of the bandpass effect of the MIRAS instrument. No instrumental noise or outliers are added to the input ground truth image. For this experiment, the outliers' support in the proposed model was set to 1 pixel, following the strategy described in Section IV-A. Results are displayed on Figure 9. In the first row, we show the result of computing $T_{L1B} = \mathbf{J}^+\mathbf{G}T_{gt}$. The error reported on the corresponding table is only related to the bandpass matrix $\mathbf{G}$ and zero padding solution with its corresponding matrix $\mathbf{J}^+$. It is clear from this example that our method truly performs a spectral extrapolation, whose effects can be easily observed both on the image edges and on the reduction of Gibbs oscillations in the hole image. These two aspects can also be observed in the associated table, on the reported values of RMSE against the original image T and the band-limited version T_w. It is clear that the zero-padding approach is better to obtain the band-limited version of the image: the RMSE is only 0.55, against 5.10 for the proposed method. Nevertheless, when we compare against the original T image, the result obtained using the proposed method is consistently better: the RMSE is 7.86 (against 9.67 for the zero-padding method), the maximum error is much lower and also the difference image presents much less structure than on the zero-padding approach. One final remark: note that in this case, when parameter μ is set considering an outlier of 1 pixel radius, the small structures are kept (observe the presence of all Baleares islands). This suggests that, when no RFIs are present in the

image, we can set much larger value for μ, leading to better results.

B. Experiments on real data

The second set of experiments was performed on real data. Experiments were performed on several snapshots and datasets. Here, for the sake of brevity, we illustrate typical results presenting three snapshots taken on march 2010. Noise variance $\sigma^2_{\Delta V_L}$ was considered to be 0.1, which is the measurement error reported by the SMOS mission. Finally, parameter μ was set to 0.2. This corresponds to a 10 pixels support. As pointed out before, since the spatial resolution of the temperature map varies spatially, the area enclosed by a fixed number of pixels depends on each location. It is straightforward to work at a fixed spatial resolution by adjusting μ following the resolution map of the bright temperature image.

Figure 10 shows the results obtained for three snapshots, where the first row corresponds to Northern Africa, the second one to Northern Europe (Denmark and England) and the third one to Central Europe (Italy). The first row is a snapshot which does not contain any RFI. The result obtained using the proposed method illustrates two of its advantages. First, it performs a better spectral extrapolation, resulting in a much sharper image than the one recovered using zero padding with Blackman apodization. Second, it performs a reduction on Gibbs effects, which is clear when comparing these results with the ones obtained with the zero padding method.

The second an third rows show the results obtained for two snapshot containing strong RFIs. Again in this case, we can see that the proposed method yields better results: it removes RFIs from the scene, while improving the spectral extrapolation and reducing the Gibbs effects, clearly visible on the Zero Padding and Blackman results.
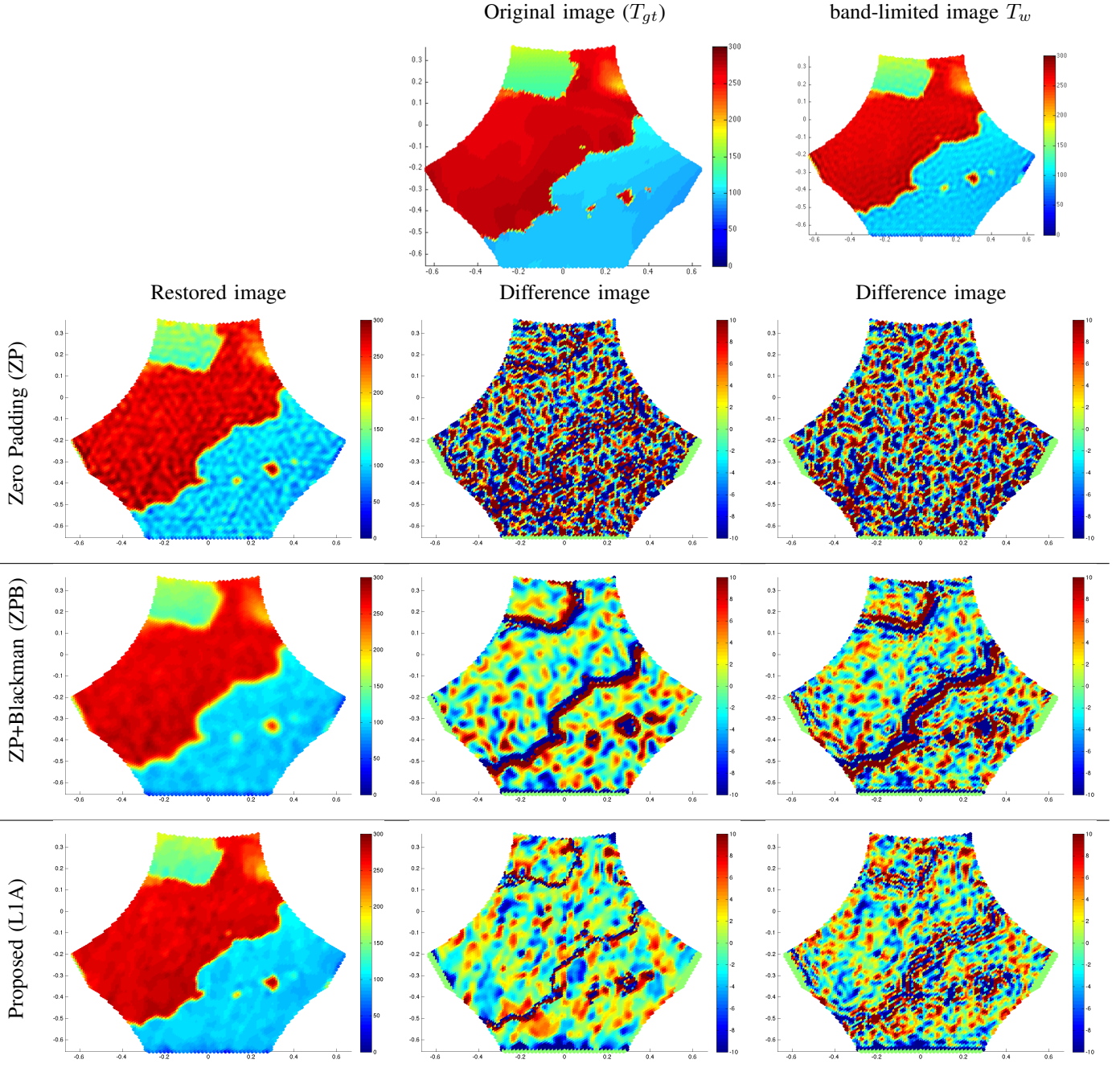
C. About computational cost

The number of iterations and convergence rate depend on each particular input data. For instance, in a regular PC with an Intel Core i7 processor and 16 GB of RAM, if no RFIs are present in the image, the method converges within one or two minutes; If RFIs are present in the image, convergence is about ten times slower (in the order of 10,000 iterations).

VI. CONCLUSIONS AND FUTURE WORK

In this work we propose a novel approach for the restoration of images acquired by the SMOS mission. Two fundamental contributions are presented: First, a variational approach that seeks to restore the image of the Earth's brightness temperature on one side, and the image of outliers or RFIs on the other side. The second contribution is the use of the visibilities or L1A data product directly for the restoration of corrupted images with RFIs, when most of the previous work use the temperatures obtained by zero padding regularisation as input (L1B data product). If we consider that the expected noise is known, the Lagrange multiplier λ is adjusted automatically by the optimization procedure considering the given noise variance, as explained in Section II-A. The method has only
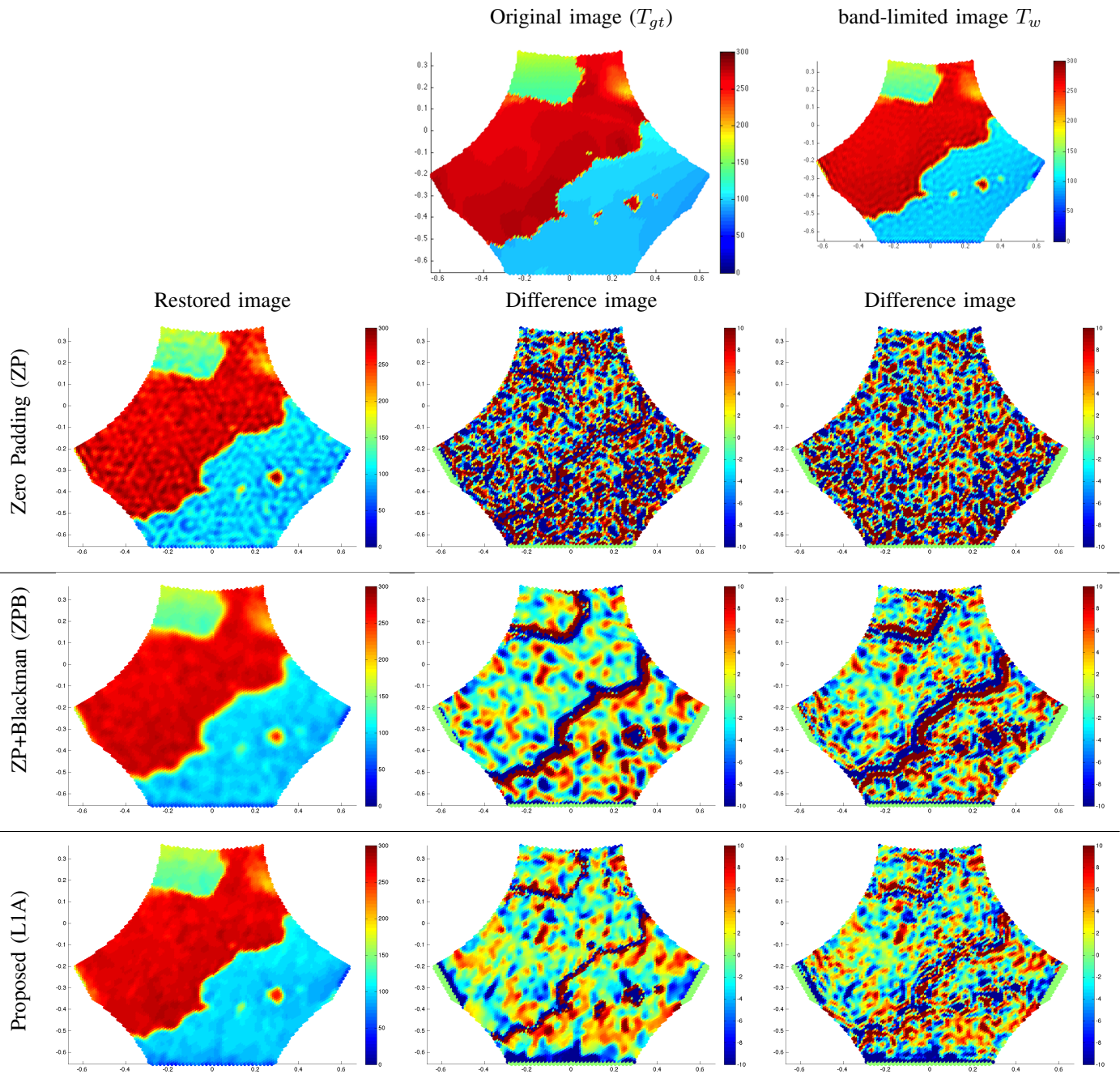
Original image (T_{gt}) band-limited image T_w

Restored image   Difference image   Difference image

	RMSE(T_{gt})	Max Error(T_{gt})	RMSE(T_w)	Max Error(T_w)
T_{ZP}	13.160	108.523	8.936	37.513
T_{ZPB}	13.612	143.147	9.933	70.018
T_{L1A}	9.2245	145.523	6.661	44.366

Fig. 5. Comparison with previous approaches when no outliers are present in the image. Error measures are expressed in Kelvin units and the error range on the images is set between [-10,10] K. For this experiment, parameter μ was set to 2.0 (corresponding to RFIs of at least 1 pixel radius). T_{ZP}, T_{ZPB} and T_{L1A} denotes the results obtained by the Zero Padding method, the Blackman apodization method, and the method proposed in this work respectively.

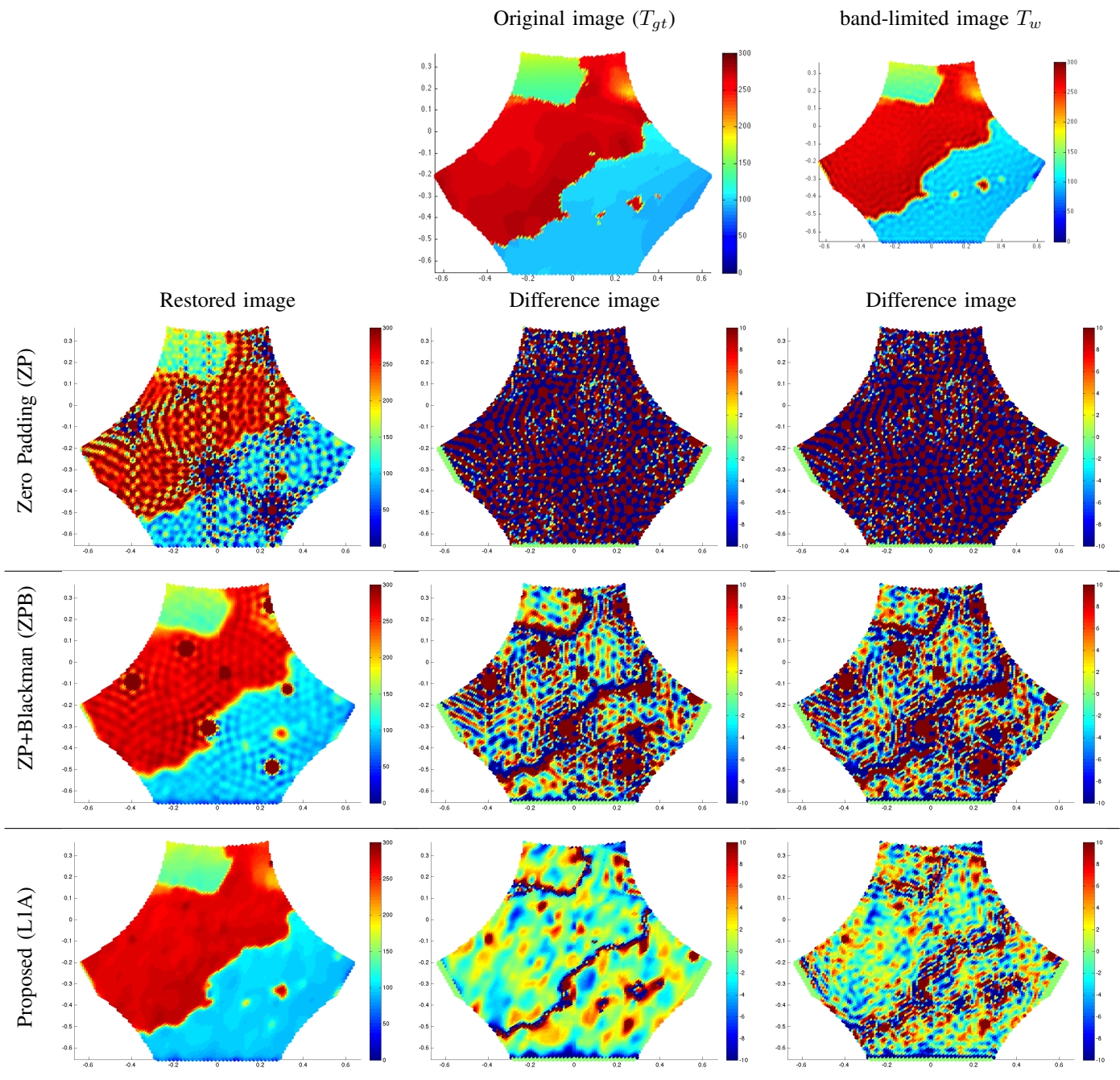|	RMSE(T_{gt})	Max Error(T_{gt})	RMSE(T_w)	Max Error(T_w)
No outliers present in the image	10.743735	142.671738	7.644904	75.479853
Outliers with values on the image range	10.416675	147.356192	7.266313	57.133400
Outliers in the range [0,35000] K	13.082794	170.410338	10.678159	149.955697
Outliers outside grid position (25000 K)	11.945941	166.184179	9.403918	133.394016

TABLE I

QUANTITATIVE RESULTS OBTAINED WITH THE PROPOSED METHOD FROM SIMULATED DATA, IN DIFFERENT SCENARIOS: WITHOUT OUTLIERS, WITH OUTLIERS INTENSITIES WITHIN THE RANGE OF THE IMAGE, WITH OUTLIERS LOCATED ON THE SAMPLING GRID POSITIONS AND OUTSIDE OF IT.
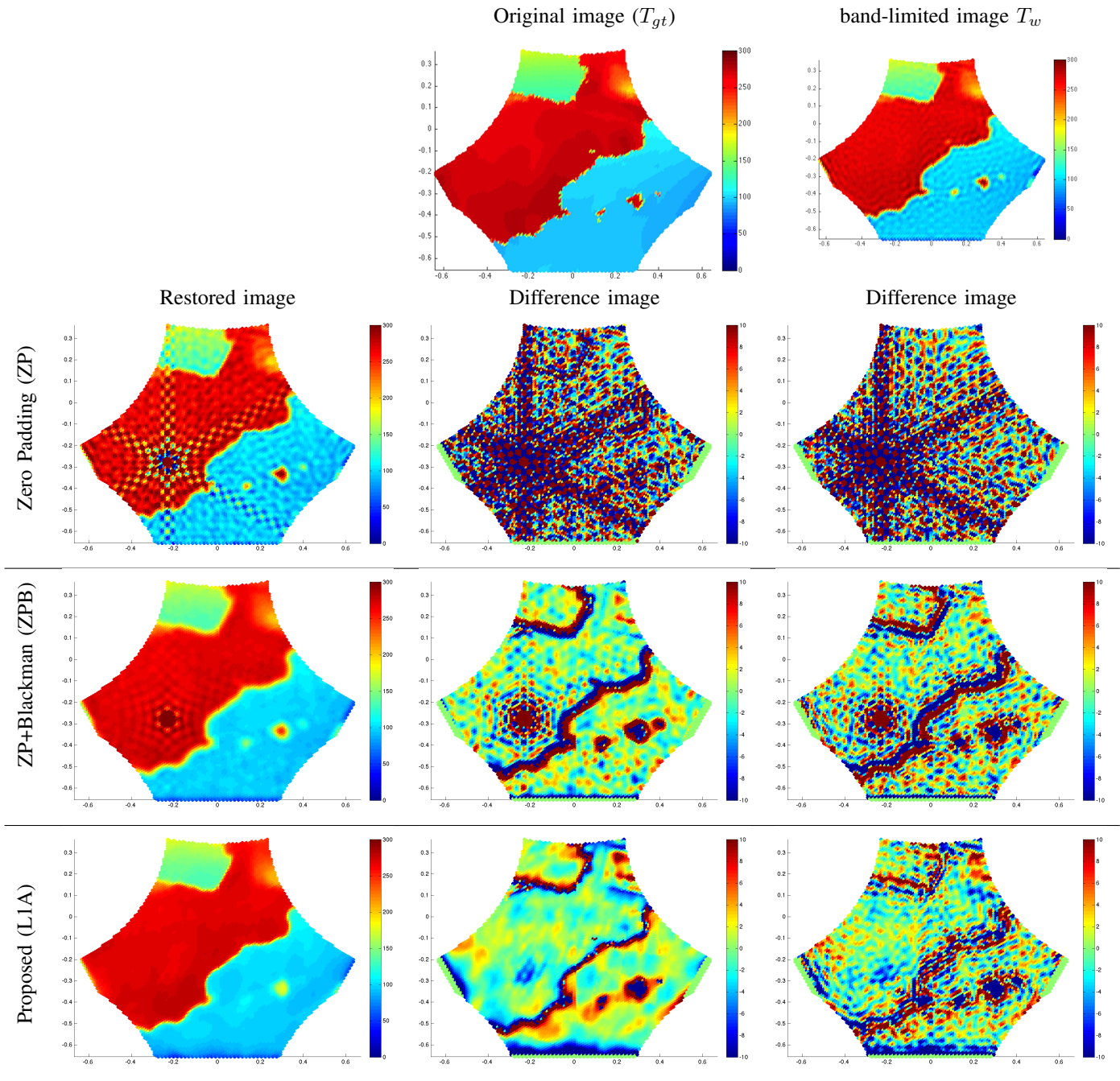
Original image (T_{gt})

band-limited image T_w

Restored image

Difference image

Difference image

Zero Padding (ZP)

ZP+Blackman (ZPB)

Proposed (L1A)

	RMSE(T_{gt})	Max Error(T_{gt})	RMSE(T_w)	Max Error(T_w)
T_{ZP}	13.733980	101.325733	9.787204	63.695435
T_{ZPB}	13.726109	140.964082	10.078270	69.381071
T_{L1A}	10.416675	147.356192	9.454039	112.115255

Fig. 6. Analysis of the different methods when the outliers intensities fall within the range of the Earth's brightness temperatures. Error measures are expressed in Kelvin. T_{ZP}, T_{ZPB} and T_{L1A} denotes the results obtained by the Zero Padding method, the Blackman apodization method, and the method proposed in this work (with $\mu = 0.2$) respectively.

Fig. 7. Results obtained with simulated data when several outliers with different intensities are present on the image. In this example, outliers are generated on the following grid positions with its corresponding values: (-0.0357,-0.3093), 35.000 K; (-0.1429,0.0619), 10000 K; (-0.3929,-0.0928), 25000 K; (0.5714, -0.1753), 800 K; (0.3304,-0.1289), 8000 K; (0.2589, 0.2629), 35000 K; (0.2589,-0.4897), 30000 K; (0.0357,-0.0515), 2000 K (Recall that this information is not used during the restoration process: it is only provided for a better understanding of the Figure.). Error measures are expressed in Kelvin units and the error range on the images is set between [-10,10] K. T_{ZP}, T_{ZPB} and T_{L1A} denotes the results obtained by the Zero Padding method, the Blackman apodization method, and the method proposed in this work (with $\mu = 0.2$) respectively.

Fig. 8. Analysis of the different methods when the the outliers are not located on the sampling grid. Note that the method performs well, with a performance similar to the one obtained when outliers are present on grid positions. In this experiment, the outlier introduced in the image has an intensity of 20000 K. Error measures are expressed in Kelvin units and the error range on the images is set between [-10,10] K. T_{ZP}, T_{ZPB} and T_{L1A} denotes the results obtained by the Zero Padding method, the Blackman apodization method, and the method proposed in this work (with $\mu = 0.2$) respectively.

Fig. 9. Comparison of extrapolation results. Error measures are expressed in Kelvin units and the error range on the images is set between [-10,10] K. T_{ZP}, T_{ZPB} and T_{L1A} denotes the results obtained by the Zero Padding method, the Blackman apodization method, and the method proposed in this work (with $\mu = 2$) respectively. See text for details.
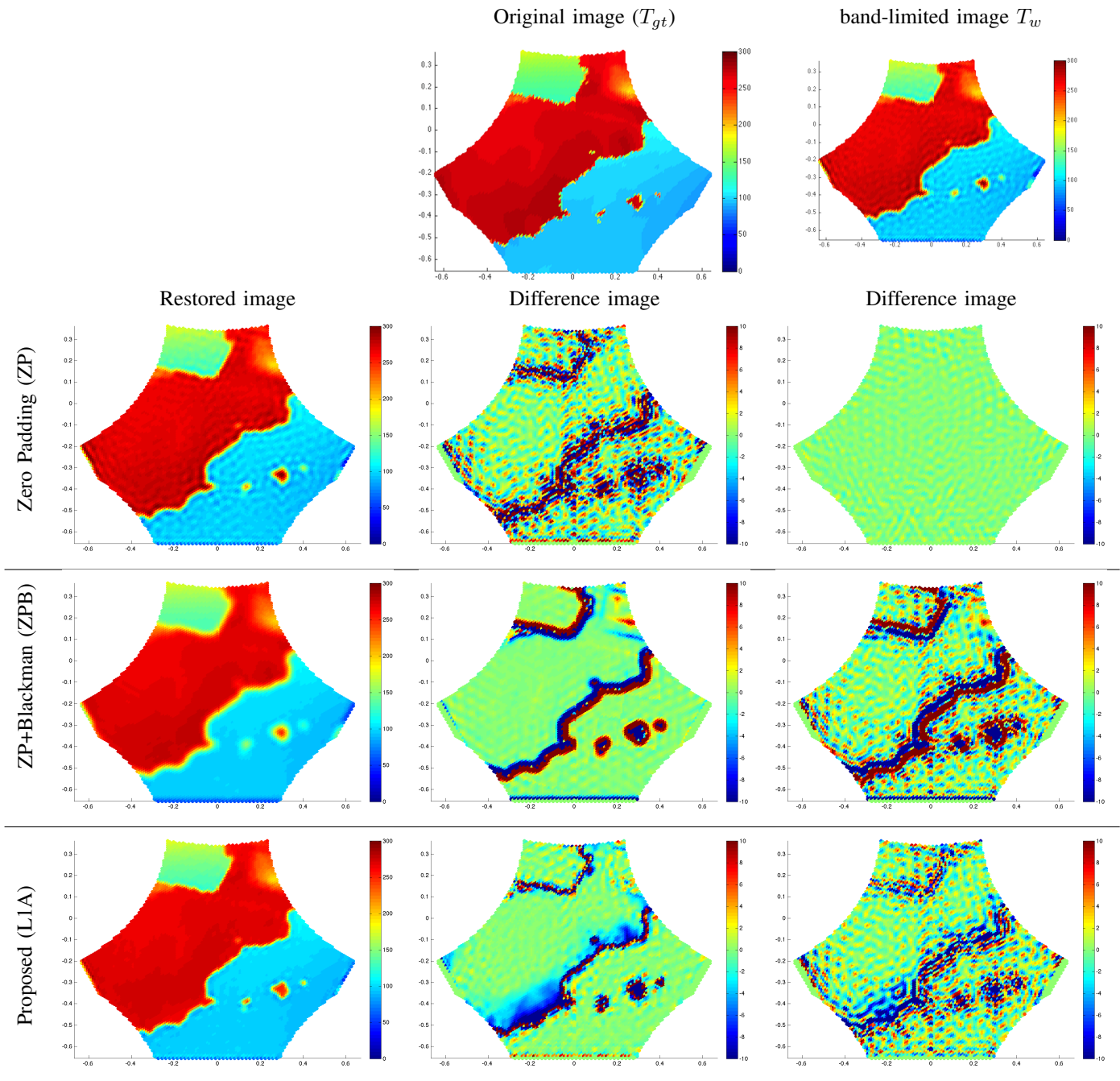
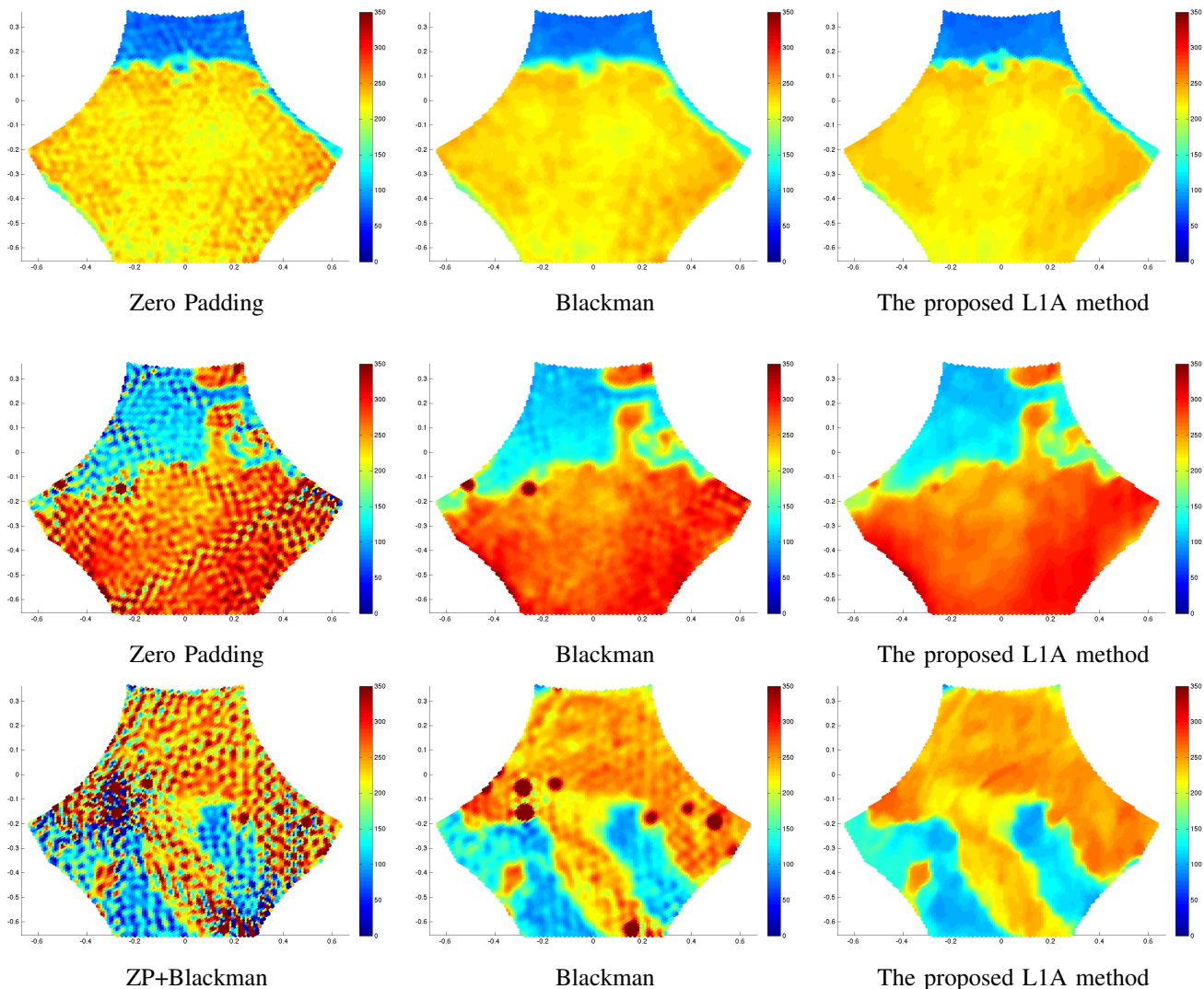|	RMSE(T_{gt})	Max Error(T_{gt})	RMSE(T_w)	Max Error(T_w)
T_{ZP}	9.671122	110.410052	0.557796	2.717416
T_{ZPB}	13.071849	144.637200	9.179310	63.214112
T_{L1A}	7.858867	89.453253	5.101420	54.833780

Fig. 10. Comparison between previous works and our method. The first row correspond to a snapshot from Northern Africa, where no outliers are present in the scene. Second and third row correspond to Northern Europe and Central Europe (With Italy clearly visible) respectively. All snapshots where acquired on march 2010. Color scale ranges from 0 to 350 Kelvin and all of the results are shown on the Extended Free of Alias Field of View (E-AF-FOV).

one free parameter μ that, as discussed in Section IV-A, controls which impulsive structures are RFIs or belong to the Earth's temperature map. Parameter μ is therefore a scale parameter that can be easily chosen based on geometric considerations. As pointed out before, scale parameters or resolution parameters are unavoidable in any image restoration problem. In most restoration techniques, the low-pass filters that are considered impose a resolution limit that fade out or even eliminate small relevant structures.

In summary, experiments on both synthetic and real data data show the ability of our method to recover the Earth's brightness temperature with high precision, and confirm the suitability of our method, outperforming previous approaches proposed for SMOS images restoration. On the other hand, the drawback of the proposed method is the time it requires to restore each snapshot (several minutes when RFIs are present in the scene) due to its iterative nature and to the fact that in each iteration an inverse FFT has to be computed. In this sense, the nominal approach runs almost instantly since it requires a single FFT inversion.

APPENDIX A
PROXIMAL OPERATORS

For the sake of completeness, we include the derivation of each of the proximal operators involved in this work. We first recall the proximal operator definition of a functional E:

$$\text{prox}_{\gamma E}(x) = \arg\inf_{y} \ E(y) + \frac{1}{2\gamma}\|x - y\|^2.$$

Proposition 2 (Proximal operators of ℓ_0 and ℓ_1 norms)**.**
$\text{prox}_{\gamma\|\cdot\|_1}(o))[i] = s_\gamma(o[i]), \quad \text{prox}_{\gamma\|\cdot\|_0}(o))[i] = \tau_\gamma(o[i]),$
where

$$s_\gamma(t) = \begin{cases} \text{sign}(t)\,(|t| - \gamma) & \text{if } |t| \geq \gamma \\ 0 & \text{if } |t| < \gamma \end{cases}$$

$$\tau_\gamma(t) = \begin{cases} t & \text{if } |t| \geq \gamma \\ 0 & \text{if } |t| < \gamma \end{cases}$$

are the soft *and* hard thresholding *operators, respectively.*

A. *Derivation of the proximal operator of $TV_{\mathcal{H}}$*

We recall first the definition of the total variation semi-norm of an image $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$:

$$J(u) = \sup_{\psi \in C^1(\Omega, \mathbb{R}^2), \|\psi\|_{L^\infty(\Omega)} \leq 1} \left\{ \int_\Omega u(x) \operatorname{div} \psi(x) dx \right\},$$

where $C^1(\Omega, \mathbb{R}^2)$ denotes the space of continuously differentiable functions of compact support contained in Ω. We also recall the discrete formulation, considering 2D images of size $N \times N$ ($u \in X = \mathbb{R}^{N \times N}$):

$$\mathrm{TV}(u) = \sum_{1 \leq i,j \leq N} |(\nabla u)_{i,j}|$$

and on its more general form:

$$\mathrm{TV}(u) = \sup_{p} <p, \nabla u>_Y$$

with $|p| \leq 1$ and $Y = X \times X$. With this notation, $\mathrm{prox}_{\gamma \mathrm{TV}}$ can be expressed:

$$\mathrm{prox}_{\gamma \mathrm{TV}}(u) = \arg \inf_{y} \ \mathrm{TV}(y) + \frac{1}{2\gamma} \|u - y\|^2.$$

In Section 3 [23] derives a solution of this problem. Indeed, the Euler equation for the previous equation is:

$$u - y + \gamma \partial \, \mathrm{TV}(u) \ni 0$$

which is equivalent of:

$$u \in \partial \, \mathrm{TV}^*((y - u)/\gamma)$$

where TV^* is the dual function of TV(also known as the Legendre-Fenchel transform):

$$\mathrm{TV}^*(v) = \sup_{u} <u, v>_X - \mathrm{TV}(u)$$

writing

$$\frac{y}{\gamma} \in \frac{y-u}{\gamma} + \frac{1}{\gamma} \partial \, \mathrm{TV}^*(\frac{y-u}{\gamma})$$

we can see that $w = (y - u)/\gamma$ is the minimizer of:

$$\frac{\|w - (y/\gamma)\|^2}{2} + \frac{1}{\gamma} \mathrm{TV}^*(w)$$

Because TV^* is known to be the characteristic function of some closed convex set K, we deduce that

$$w = \Pi_K(y/\gamma)$$

where Π_K denotes the nonlinear projection to K.
This derivation is in fact independent of the ∇ operator on TV. In fact, we can generalize the previous result to a more general form. If we redefine

$$\mathrm{TV}(u) = \sup_{p} <p, Au>_Y$$

with $|p| \leq 1$ for any linear operator A, we obtain the same result, with the only difference that set K is defined accordingly:

$$K = \{A^*v : v \in C_c^1(\Omega, \mathbb{R}^2), \|v\|_\infty \leq 1\}$$

where A^* is the adjoint operator. Finally, projection $\Pi_K(y)$ can be obtained by solving:

$$\min \|\gamma A^* p - y\|^2$$

with $\|p\| \leq 1$. In the discrete formulation, and again following [23], the existence of the KKT conditions yields the existence of lagrange multipliers $\alpha_{i,j}$, resulting:

$$A(\gamma A^* p - y)_{i,j} - \alpha_{i,j} p_{i,j} = 0$$

for all $0 \leq i, j \leq N$. With this new formulation, the final iterative algorithm derived by [23] can be reformulated:

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau (AA^* p^n - d)_{i,j}}{1 + \tau |AA^* p^n - d|_{i,j}}$$

with $d = A(y/\gamma)$. For the specific case where $A = \operatorname{div}$, [23] has proven (Theorem 3.1) that, if $0 < \tau \leq \frac{1}{\kappa^2}$ then $\gamma \operatorname{div} p^n$ converges to $\Pi_{\gamma K}(y)$ as $n \to \infty$, where κ is the norm of operator div. Nevertheless, the demonstration is generic and we can easily extend the results to the more general form where A is any linear operator. Thus, we have the following

Proposition 3. *Let $\tau \leq 1/\kappa^2$ with κ the norm of operator A^*. Then $\gamma A^* p^n$ converges to $\Pi_{\gamma K}(y)$ as $n \to \infty$.*

In our case, the spectral TV minimization consists in restricting the frequency domain to the cell $\mathcal{H}$, *i.e.*

$$\mathrm{prox}_{\gamma \, \mathrm{TV}_{\mathcal{H}}}(u) = \arg \inf_{u \in BL(\mathcal{H})} \mathrm{TV}(u) + \frac{1}{2\gamma} \|u - y\|^2.$$

where $BL(\mathcal{H})$ denotes the space of band limited functions with spectral support within $\mathcal{H}$. This constraint can be integrated into the TV operator by means of the indicator function:

$$\iota_{BL(\mathcal{H})}(x) = \left\{ \begin{array}{ll} 0 & \text{if } x \in BL(\mathcal{H}) \\ +\infty & \text{if } x \notin BL(\mathcal{H}) \end{array} \right.$$

Then the proximal operator for the spectral TV becomes

$$\mathrm{prox}_{\gamma \, \mathrm{TV}_{\mathcal{H}}}(u) := \mathrm{prox}_{\gamma TV + \iota_{BL(\mathcal{H})}}(u).$$

This can finally be expressed:

$$\arg \inf_{u \in BL(\mathcal{H})} \sum_{0 \leq i,j \leq N} |\nabla F^{-1} Z F(u))_{i,j}| + \frac{1}{2\gamma} \|u - y\|^2.$$

where Z is a mask operator that restrict the domain to be contained on $\mathcal{H}$. Going back to Theorem 3, if we note $A = \nabla F^{-1} Z F$, the adjoint operator is

$$A^* = F^{-1} Z^* F \operatorname{div}$$

where Z^* is the conjugate of Z, which has a simple form since Z is in fact a mask (1 for points on $\mathcal{H}$ and 0 otherwise). To assure convergence, we finally set $\tau < 1/\kappa^2$ where $\kappa = \sup_{\|p\| \leq 1} \|F^{-1} Z^* F \operatorname{div} p\|$ can be determined empirically.

ACKNOWLEDGMENT

REFERENCES

[1] Y. Kerr, P. Waldteufel, J.-P. Wigneron, S. Delwart, F. Cabot, J. Boutin, M.-J. Escorihuela, J. Font, N. Reul, C. Gruhier, S. Juglea, M. Drinkwater, A. Hahne, M. Martin-Neira, and S. Mecklenburg, "The SMOS mission: New tool for monitoring key elements ofthe global water cycle," *Proceedings of the IEEE*, vol. 98, no. 5, pp. 666–687, May 2010.

[2] K. D. McMullan, M. Brown, M. Martin-Neira, W. Rits, S. Ekholm, J. Marti, and J. Lemanczyk, "SMOS: The payload," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 3, pp. 594–605, March 2008.

[3] M. Martin-Neira, Y. Menard, J. Goutoule, and U. Kraft, "MIRAS, a two-dimensional aperture synthesis radiometer," in *IEEE International Geoscience and Remote Sensing Symposium 1994. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation*, vol. 3, Aug 1994, pp. 1323–1325.

[4] Y. Kerr, J. Font, P. Waldteufel, A. Camps, J. Bara, I. Corbella, F. Torres, N. Duffo, M. Vallilossera, and G. Caudal, "New radiometers: SMOS - a dual pol L-band 2D aperture synthesis radiometer," in *Aerospace Conference Proceedings, 2000 IEEE*, vol. 5, 2000, pp. 119–128.

[5] I. Corbella, N. Duffo, M. Vall-llossera, A. Camps, and F. Torres, "The visibility function in interferometric aperture synthesis radiometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 8, pp. 1677–1682, 2004.

[6] E. Anterrieu, "A resolving matrix approach for synthetic aperture imaging radiometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1649–1656, Aug 2004.

[7] D. Petersen and D. Middleton, "Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidean space," *Information and Control*, vol. 5, pp. 279–323, 1962.

[8] A. Camps, J. Bara, I. Sanahuja, and F. Torres, "The processing of hexagonally sampled signals with standard rectangular techniques: application to 2-D large aperture synthesis interferometric radiometers," *IEEE Transactions onGeoscience and Remote Sensing*, vol. 35, no. 1, pp. 183–190, Jan 1997.

[9] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy; 2nd ed.* Weinheim: Wiley-VCH, 2001.

[10] A. Camps, M. Vall-llossera, I. Corbella, N. Duffo, and F. Torres, "Improved image reconstruction algorithms for aperture synthesis radiometers," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 1, pp. 146–158, Jan 2008.

[11] R. Castro, A. Gutiérrez, and J. Barbosa, "Iterative thresholding for sparse approximations," *IEEE Transactions on geoscience and remote sensing*, vol. 50, no. 5, pp. 1440–1447, May 2012.

[12] A. Camps, M. Vall-Llossera, N. Duffo, M. Zapata, I. Corbella, F. Torres, and V. Barrena, "Sun effects in 2-D aperture synthesis radiometry imaging and their cancelation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 6, pp. 1161–1167, 2004.

[13] Y. Soldo, A. Khazaal, F. Cabot, P. Richaume, E. Anterrieu, and Y. Kerr, "Mitigation of RFIs for SMOS: A distributed approach," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 11, pp. 7470–7479, Nov 2014.

[14] R. Oliva, S. Nieto, and F. Felix-Redondo, "RFI detection algorithm: Accurate geolocation of the interfering sources in SMOS images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 10, pp. 4993–4998, Oct 2013.

[15] V. González-Gambau, A. Turiel, E. Olmedo, J. Martínez, I. Corbella, and A. Camps, "Nodal sampling: A new image reconstruction algorithm for SMOS," *IEEE Trans. Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2314–2328, 2016. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2015.2499324

[16] E. Anterrieu, P. Waldteufel, and A. Lannes, "Apodization functions for 2-D hexagonally sampled synthetic aperture imaging radiometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 12, pp. 2531–2542, Dec 2002.

[17] J. Preciozzi, P. Musé, A. Almansa, S. Durand, A. Khazaal, and B. Rougé, "SMOS images restoration: A sparsity-based variational approach," in *IEEE International Geoscience and Remote Sensing Symposium*, 2014.

[18] C. S. Ruf, C. T. Swift, A. B. Tanner, and D. M. Le Vine, "Interferometric synthetic aperture microwave radiometry for the remote sensing of the earth," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 26, no. 5, pp. 597–611, 1988.

[19] J. Freitas, J. Correia, and N. Catarino, "SMOS L1 processor L1a to L1b data processing model," DEIMOS Engenharia, Lisboa, Portugal, Tech. Rep. SO-DS-DME-L1PP-0008, July 2009.

[20] H. Uzawa, "Iterative methods for concave programming," *Studies in linear and nonlinear programming*, vol. 6, 1958.

[21] F. Malgouyres and F. Guichard, "Edge direction preserving image zooming: a mathematical and numerical analysis," *SIAM Journal on Numerical Analysis*, vol. 39, no. 1, pp. 1–37, 2001.

[22] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[23] A. Chambolle, "An algorithm for total variation minimization and application," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, 2004.

[24] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[25] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *SIAM Journal on Multiscale Modeling and Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[26] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *The Journal of Fourier Analysis and Applications*, vol. 62, pp. 291–294, January 2008.

[27] M. Nikolova, "Local strong homogeneity of a regularized estimator," *SIAM Journal on Applied Mathematics*, vol. 61, no. 2, pp. 633–658, 2000.

[28] K. Jalalzai, "Some remarks on the staircasing phenomenon in total-variation based image denoising," *ArXiv e-prints*, Feb. 2014.

[29] L. Moisan, "How to discretize the total variation of an image?" *ICIAM07 Minisymposia*, 2007.

[30] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vis.*, vol. 40, no. 1, pp. 120–145, May 2011. [Online]. Available: http://dx.doi.org/10.1007/s10851-010-0251-1

[31] R. Abergel and L. Moisan, "The Shannon Total Variation," 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01349516

[32] R. Abergel, "Several mathematical models and fast algorithms for image processing," PhD, Université Paris Descartes, 2016.

[33] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[34] ——, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[35] A. Camps, I. Corbella, M. Vall-llossera, N. Duffo, F. Marcos, F. Martínez-Fadrique, and M. Greiner, "The SMOS end-to-end performance simulator: Description and scientific applications," in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 1, 2003, pp. 13–15.