Introducing spatial regularization in SAR tomography reconstruction

Clément Rambour, Loïc Denis, Florence Tupin, Hélène Oriot

Introducing spatial regularization in SAR tomography reconstruction

Clément Rambour *Student Member, IEEE*, Loïc Denis, Florence Tupin *Senior Member, IEEE*, Hélène Oriot

Abstract—The resolution achieved by current Synthetic Aperture Radar (SAR) sensors provides detailed visualization of urban areas. Spaceborne sensors such as TerraSAR-X can be used to analyze large areas at a very high resolution. In addition, repeated passes of the satellite give access to temporal and interferometric information on the scene. Because of the complex 3-D structure of urban surfaces, scatterers located at different heights (ground, building façade, roof) produce radar echoes that often get mixed within the same radar cells. These echoes must be numerically unmixed in order to get a fine understanding of the radar images. This unmixing is at the core of SAR tomography. SAR tomography reconstruction is generally performed in two steps: (i) reconstruction of the so-called tomogram by vertical focusing, at each radar resolution cell, to extract the complex amplitudes (a 1-D processing); (ii) transformation from radar geometry to ground geometry and extraction of significant scatterers. We propose to perform the tomographic inversion directly in ground geometry in order to enforce spatial regularity in 3-D space. This inversion requires solving a large-scale non-convex optimization problem. We describe an iterative method based on variable splitting and the augmented Lagrangian technique. Spatial regularizations can easily be included in this generic scheme. We illustrate on simulated data and a TerraSAR-X tomographic dataset the potential of this approach to produce 3-D reconstructions of urban surfaces.

Index Terms—tomographic SAR inversion, dense urban areas, TerraSAR-X, Compressed Sensing, inverse problems, 3-D reconstruction

I. INTRODUCTION

SAR imagery is a powerful modality for the observation and the interpretation of natural and man-made areas. The existing diversity of SAR spaceborne sensors, with their different spatial and temporal resolutions, has given rise to numerous applications. They provide useful information for the monitoring of large areas for man-made changes, natural disasters or displacement of the ground. Moreover, depending on the operating bandwidth, the radar wave can penetrate some environments (vegetation, ice) and provide information about their structure or the presence of underneath objects [1]. Over urban areas, stacks of co-registered high resolution SAR images can be jointly processed to unmix SAR echoes from

scatterers at different 3-D locations and thus reconstruct the 3-D reflectivity of the scene. The 3-D representation of cities has many applications such as crisis management, movement monitoring or change detection.

The spatial resolution of modern sensors such as TerraSAR-X or COSMO-SKYMED can be better than one meter, which is much finer than the average height of observed buildings in city areas. At these resolutions, urban surfaces (façades, roofs) can be recovered. However, as the dense urban scenes contain a high number of back-scattering objects, scatterers from different structures are projected within the same radar resolution-cell. The actual tridimensional location of these scatterers can be reconstructed by combining images from repeated passes with slightly different angles of observation, by SAR tomography. SAR tomography can be seen as an extension of the SAR focusing to 3 dimensions, the separation of scatterers projected within a given radar cell being performed thanks to a synthetic aperture formed by the different trajectories of the SAR antenna in the tomographic stack [2], [3].

In the ideal case of equi-spaced trajectories, focusing in the direction orthogonal to the line of sight can be simply performed by application of the inverse discrete Fourier transform. The resolution of this focusing is inversely proportional to the maximal orthogonal baseline. Increasing the number of tracks within this maximal orthogonal baseline improves the sampling, hence reduces height ambiguities. However, the vertical resolution is generally much worse compared to the resolution in azimuth and range directions. Moreover, the baselines are generally irregularly distributed which produces side-lobes higher than expected and degrades the interpretation of the reconstructed volume. Several spectral super-resolution techniques have been introduced to overcome these phenomena [4], [5], leading to satisfying results on homogeneous areas dominated by volumic scattering such as forests or landscape areas [1], [6]. These estimators rely on the analysis of the covariance matrix. The estimation of this matrix is not a trivial task because it involves some form of spatial averaging [7] and becomes very challenging in dense urban areas due to the high spatial heterogeneity [8], [9]. Rather than exploiting the covariance matrix, some works on very high resolution data use directly the time domain backprojection of the signal [10]. When backscattering elements are discrete rather than continuous, the well known Compressed Sensing (CS) approach can produce super-resolved tomographic reconstruction by performing the backprojection of the data under sparsity constraints [11], [12]. Since this approach does not require the estimation of a covariance matrix, it involves no spatial

C. Rambour is with the LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France, and also with the ONERA, The French Aerospace Laboratory, 91761 Palaiseau France (e-mail: clement.rambour@telecom-paristech.fr).

L. Denis is with the Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France (e-mail: loic.denis@univ-st-etienne.fr).

F. Tupin is with the LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France (e-mail: florence.tupin@telecom-paristech.fr).

H. Oriot is with ONERA, The French Aerospace Laboratory, 91761 Palaiseau, France (e-mail: helene.oriot@onera.fr).

filtering and is more effective at reconstructing multiple close-by point scatterers.

CS appears to be well suited to the inversion of stacks of tomographic SAR images obtained over dense urban areas. It can however be criticized for its lack of spatial regularity: pixels are processed independently while many urban structures are organized in lines or planes. Including a spatial regularization can be expected to improve the robustness to noise, reduce the ambiguities and better preserve the signal coming from the buildings. Moreover, the sparse estimation step of the signal is generally hard to tune. Due to the high dynamic of SAR images, it often appears that either weak parts of the signal are suppressed during the estimation, or outliers are detected. To refine the results, the volume reconstructed with CS is generally post-processed by estimating the number of scatterers in order to select only the most significant points, see [13], [14]. The result is then a set of discrete points defined by their 3-D localization and complex reflectivity. As the distribution of the reflectivity in dense urban configurations is hard to model, this approach often fails to select points with low reflectivities. Close to CS, the recent FAST-SUP-GLRT detector [15] avoids post processing selection by applying a sub-optimal statistical test taking into account the distribution of the data based on an approximated ℓ_0 norm minimization. Even if it doesn't take into account the geometry of the scene, the structure of the data can be more accurately represented than with conventional CS.

A recent strategy to enhance the discrete reconstructions obtained with CS consists of post-processing the SAR tomographic (post-processed) points clouds. It is mainly based on shape detection and surface fitting using priors on the urban geometry. It has been shown to recover some of the information lost during point selection and even to fill missing parts of the scene [16] [17]. Those approaches give satisfying geometrical 3-D modeling of urban environments but lack from a direct connection to the data. In this paper, we consider an alternative approach that enforces spatial regularity during the tomographic inversion step. Since the output of our method is a sparse volume, a discrete set of points can be extracted and any pipeline already designed to process tomographic point clouds can possibly be applied as a post-processing.

In this paper, we present a new algorithm that performs the tomographic 3-D inversion of a stack of SAR images combining sparse priors as in CS and geometrical ones using a ground coordinate back-projection operator as in [10]. To derive our algorithm, we account for two inherent characteristics of SAR imaging: the geometrical distortions induced by the geometry of acquisition and the use of Single Look Complex (SLC) data. First, we present in details the concept of SAR tomography and the used notations (section II). Then, we address the problem of using priors corresponding to the natural properties of the urban environment. The third section is dedicated to solving the tomographic inverse problem, which is a non-convex and high-dimensional optimization problem (section III). In the fourth section, we illustrate that our algorithm performs well both on synthetic data with controlled sensor parameters, scene composition and geometry and on real data (section IV). The quality of our results is evaluated using the

accuracy and completeness defined in [18] [19]. We use a stack of 40 TerraSAR-X images of the city of Paris to test our algorithm on real data. The results are evaluated on some buildings using a 3-D modeling of the scene given by the French Geographic Institute (IGN) as ground truth.

II. SAR TOMOGRAPHY

A. Forward model for SAR tomography in radar geometry

A SAR tomographic stack consists of N SAR Single Look Complex (SLC) images acquired in interferometric configuration. Each SAR image of the stack corresponds to a different trajectory of the sensor over the scene. We consider all images to have been co-registered with respect to a master image in a preprocessing step. To simplify the notations, in the following this master image will correspond to the first image of the stack. The geometry of a simple 3-D scene and of the sensor trajectories is illustrated in Fig. 1.

The complex value $\underline{v}_n(x,r)$ obtained at pixel coordinates (x,r) after SAR synthesis of the n-th image corresponds the convolution of the complex 3-D reflectivity with the Point Spread Function (PSF) of the sensor [2], [20]:

$$
\underline{v}_n(x,r) = \iiint \underline{f}(x - x', r - \rho_{n;y,z})\underline{u}(x,y,z).
$$
$$
\exp\left(-\frac{4j\pi}{\lambda}\rho_{n;y,z} + j\varphi_{\text{atmo}}\right)\mathrm{d}x'\mathrm{d}y\mathrm{d}z + \underline{\epsilon} \quad (1)
$$

To better identify real-valued and complex-valued variables in the description of our algorithm (section III), we underline all complex-valued variables throughout the paper. Also the x and azimuth axis coincide to avoid complicated notations without loss of generality. Here x and r stand for the discrete coordinates of the focused data, $\rho_{n;y,z}$ is the distance between the nth sensor and a scatterer at a position (y,z) for a given x. φ_{atmo} is the phase shift corresponding to the Atmospheric Phase Screen (APS). $\underline{f}$ corresponds to the PSF and depends of the sensors and the processing of the data. $\underline{\epsilon}$ is a white additive Gaussian noise.

In this paper, we will use two classical approximations in SAR tomography. Firstly, the PSF is substituted with a boxcar function that may have a near zero width depending on the knowledge of the data. Then, to be in the same framework as most of the other state of the art tomographic approaches, we use the interferometric images with respect to the master image as the input tomographic stack. The phase difference is then approximated as a linear function of the elevation of the scatterers [20]. To remove the APS contribution we use the method detailed by Ferretti and Al. [21]. Without assuming movement of the scatterers, the equation (1) becomes the sum of the complex amplitudes scattered by all objects that project within the resolution cell [11], [12], [20]

$$
\underline{v}_n(x,r) = \int_{(y,z)\in\Delta_r} \underline{u}(x,y,z)\exp\bigl(-j\xi_n h(y,z)\bigr)\mathrm{d}y\mathrm{d}z + \underline{\epsilon}
$$
$$(2)$$

The r-th radar resolution cell is defined by $\Delta_r = \{(y,z)\,|\,r - \delta_r/2 \le \rho_{1;y,z} \le r + \delta_r/2 \text{ and } z \le z_{\max}(y)\}$ with δ_r the step
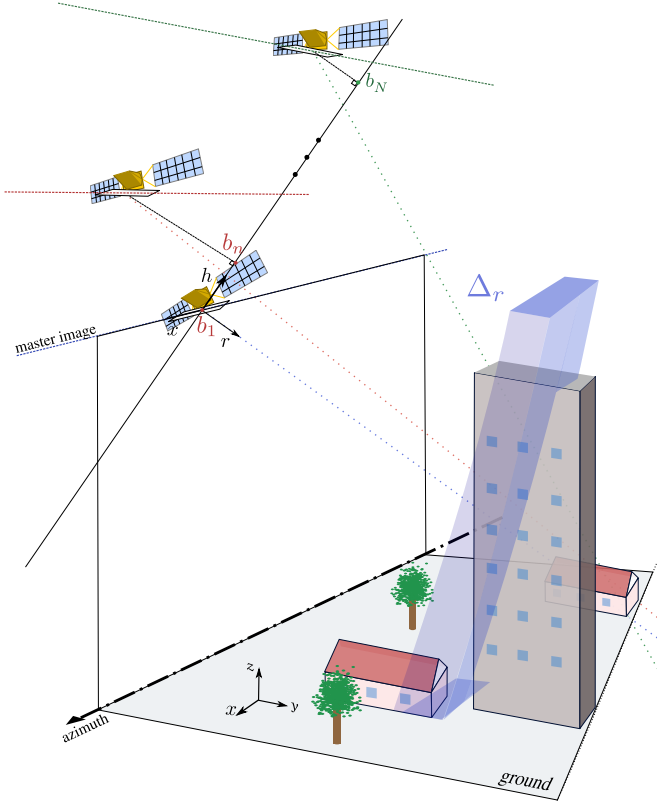
Fig. 1. Configuration of the tomographic acquisition of a 3-D scene. The scene is naturally described by the coordinate system (x, y, z) whereas the data are described in the sensor coordinate system (x, r, h).

in range and $z_{\max}(y)$ the maximum height for a point at (x, y) location to be illuminated by the radar. The chosen approximations allow us to use classical algorithms to compensate the APS and to reduce the size of the operators we will then define to perform the inversion. Here as we are also interested in a comparison with other SAR tomographic estimators we chose to keep this classical interferometric framework. Moreover, using directly the equation (1) can also be done but the phase calibration may be harder.

We denote by x, y and z the coordinates of a reflector in 3-D space and by x, r and h its coordinates in radar space. In equation (2), Δ_r is the extension of the radar resolution cell along the elevation direction h, $\underline{u}(x, y, z)$ is the complex reflectivity of the scatterer located at position (x, y, z), and $h(y, z)$ its elevation. The parameter $\xi_n = \frac{4\pi b_n}{\lambda r}$ is the spatial frequency associated to the sampling of the scene for each baseline. b_n is the baseline n and λ the radar wavelength. Random variable $\underline{\epsilon}$ models the complex additive white noise due to thermal noise.

Consider a given radar resolution cell. We will drop in the following the indices identifying the spatial location of that cell. We form the vector $\underline{u} = \left(\underline{u}_1 \cdots \underline{u}_D\right)^T$ that collects the complex reflectivities of all scatterers seen within the same radar resolution cell and $\boldsymbol{h} = \left(h_1 \cdots h_D\right)^T$ their elevation. After discretization of (2), we can express $\underline{v} \in \mathbb{C}^N$, the collection of SLC values in the tomographic stack at the pixel of interest, as a linear combination of the complex signal

backscattered by each of the D scatterers:

$$\underline{v} = \underline{\boldsymbol{A}}(\boldsymbol{h})\underline{u} + \underline{n} \,, \tag{3}$$

where $\underline{\boldsymbol{A}}(\boldsymbol{h}) \in \mathbb{C}^{N \times D}$ is the steering matrix. The d-th column $[\underline{\boldsymbol{A}}(\boldsymbol{h})]_d$ of $\underline{\boldsymbol{A}}(\boldsymbol{h})$ corresponds to the steering vector $\underline{\boldsymbol{a}}(h_d)$ associated to the elevation of the d-th scatterer:

$$[\underline{\boldsymbol{A}}(\boldsymbol{h})]_d = \underline{\boldsymbol{a}}(h_d) = \left(\exp(-\underline{j}\xi_1 h_d) \quad \cdots \quad \exp(-\underline{j}\xi_N h_d)\right)^T \tag{4}$$

B. Spectral Analysis in SAR Tomography

Many strategies exist to invert (3) in order to estimate the heights $\boldsymbol{h}$ and reflectivities $\underline{u}$. Most of them come from the Direction Of Arrival (DOA) literature and have been successfully used for SAR tomography on homogenous scenes such as forest or ice. Non-parametric methods such as beamforming or Capon filters are fast to implement and offer a good global representation of the reconstructed scene [4], [5], [22], [23]. Parametric methods achieve super resolution by exploiting priors on the estimated data. In many cases, a limited number of point-like scatterers is assumed. When considering a given number of such scatterers, the Maximum Likelihood (ML) estimator can be applied to estimate the location and reflectivity of each scatterer. Exact computation of the ML estimator requires a prohibitive computation time when more than 2 or 3 scatterers are considered, because of the combinatorial nature of the optimization problem. Suboptimal methods are therefore generally used specially when dealing with big images. MUSIC [24] is one of the most popular parametric estimators used to estimate the position of a finite number of sources. The localization of the D sources is obtained by finding the steering vectors with minimal norm in the eigen-subspace of $\underline{\boldsymbol{R}}$ associated to the $(N - D)$ smaller eigenvalues. Since this approach is adapted, by construction, to separate discrete scatterers, it is not well suited to the reconstruction of continuous reflectivity profiles. In urban areas, it offers fair performances when the reflectivity profiles can be accurately represented by sparse distributions. The pseudo-spectrum given by the inverse of the projection is also sometime used to analyze even continuous tomographic profiles [25] [26]. Very similar to MUSIC, but derived as an approximation of ML, Weighting Subspace Fitting (WSF) [27] methods are designed to achieve an even better resolution than MUSIC. By using both the signal and the noise subspace, it has been proven that WSF can be used to separate the distribution of the leaves in a forest and an artificial target underneath [1]. All those methods require to estimate the covariance matrix at each pixel, which can be a very hard task in heterogeneous regions such as dense urban areas. Since it takes at least N^2 pixels to get a full-rank sample covariance matrix, the more images in the tomographic stack, the strongest the spatial filtering necessary to estimate the covariance matrices (hence, the strongest the range and azimuth resolution loss).

C. Compressed Sensing in SAR Tomography

The use of Compressed Sensing (CS) for SAR tomography is recent. It is well suited to the urban environment [28]

as it does not require to estimate $\underline{\boldsymbol{R}}$. The estimation of the reflectivities profile $\underline{\boldsymbol{u}}$ along the height direction, for a given SAR resolution cell, is obtained by solving the following optimization problem:

$$\min_{\underline{\boldsymbol{u}}} ||\underline{\boldsymbol{u}}||_0 \quad \text{s.t.} \quad \underline{\boldsymbol{v}} = \underline{\boldsymbol{A}}(\boldsymbol{h}_0)\,\underline{\boldsymbol{u}} \tag{5}$$

As the vector of heights $\boldsymbol{h}$ is one of the unknown, the elevation axis is sampled into D_0 bins forming the array $\boldsymbol{h}_0$. The solution to the combinatorial problem (5) can be approximated using the classical convex relaxation of the ℓ_0 pseudo-norm into an ℓ_1 norm:

$$\hat{\underline{\boldsymbol{u}}} = \arg\min_{\underline{\boldsymbol{u}}} \frac{1}{2}||\underline{\boldsymbol{A}}(\boldsymbol{h}_0)\underline{\boldsymbol{u}} - \underline{\boldsymbol{v}}||_2^2 + \mu_1||\underline{\boldsymbol{u}}||_1 \tag{6}$$

where μ_1 is a Lagrange multiplier. Despite its very good performances on urban areas, this approach has some drawbacks. Indeed, the matrix $\underline{\boldsymbol{A}}(\boldsymbol{h}_0)$ does not in general guarantee to satisfy either the Restrictive Isometry Property (RIP) or an incoherence condition [29] required to obtain reconstruction guarantees when using the relaxed problem (6). When choosing a large number of samples D_0 in the height direction in order to achieve a super-resolution, these conditions are largely violated. RIP essentially indicates that any random sub-matrix of $\underline{\boldsymbol{A}}(\boldsymbol{h}_0)$ composed from D random columns of $\underline{\boldsymbol{A}}(\boldsymbol{h}_0)$ is nearly orthogonal which preserves the energy of any D-sparse vector $\underline{\boldsymbol{u}}$. The incoherence insures that a highly localized signal gets spread over all the columns of $\underline{\boldsymbol{A}}(\boldsymbol{h}_0)$ which makes the reconstruction robust to random under-sampling. The consequences of the violation of RIP and incoherency are the creation of outliers and/or multiple detections of the same scatterer in adjacent positions.

In order to refine the results obtained via a sparse reconstruction, one can include spatial priors that describe the urban geometry. As illustrated in figure 1, when performing a tomographic inversion in the radar geometry, natural geometrical properties (e.g., predominance of vertical and horizontal structures and of straight angles) are not straightforward to express. Rather than using the radar frame (r, h), we suggest in this paper to perform the tomographic inversion in the ground frame (x, y, z) in order to include structural information in the form of priors (i.e., regularization terms).

III. PROPOSED APPROACH: TOMOGRAPHIC INVERSION WITH SPATIAL REGULARIZATIONS

A. Forward model for SAR tomography in ground geometry

In order to enforce some spatial smoothness, the tomographic inversion has to be performed globally in ground coordinates. Rather than considering the collection of measurements at a given radar pixel, from now on the notations $\underline{\boldsymbol{u}} \in \mathbb{C}^{N_x.N_y.N_z}$ and $\underline{\boldsymbol{v}} \in \mathbb{C}^{N_x.N_r.N}$ will respectively refer to column vectors obtained by stacking all the values in the 3-D volume, and all the values in the tomographic stack of SAR images. N_x, N_y and N_z are the number of voxels in each direction in ground geometry while N_r is the size of SAR images in the range direction and N is the number of images in the tomographic stack. The linear operator $\underline{\boldsymbol{\Phi}} \in \mathbb{C}^{(N_x.N_r.N) \times (N_x.N_y.N_z)}$ maps the volume of complex

reflectivities in 3-D space to the complex amplitudes in the tomographic stack of SAR images, see Fig 2. Since we chose to align the x axis and the azimuth, an element of $\underline{\boldsymbol{\Phi}}$ is defined as:

$$\underline{\boldsymbol{\Phi}}_{i,j} = \begin{cases} \exp(-j\varphi) & \text{if } x_i = x_j \text{ and} \\ & r_i - \frac{\delta_r}{2} < \rho_{1;y_j,z_j} < r_i + \frac{\delta_r}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

with $\varphi = \xi_i h(y_j, z_j)$ the phase shift due to the path between voxel j with coordinates (x_j, y_j, z_j) and antenna i. As in the previous equations, $\rho_{1;y_j,z_j}$ corresponds to the distance between antenna 1 (of the master image) and the point with ground coordinates (y_j, z_j). The size of a radar pixel in the range direction is noted δ_r.

The construction of the matrix $\underline{\boldsymbol{\Phi}}$ is illustrated on Fig. 3. Matrix $\underline{\boldsymbol{\Phi}}$ is sparse: only a few entries are different from zero, so that products of the form $\underline{\boldsymbol{\Phi}}\boldsymbol{u}$ can be computed efficiently.

The observed SAR tomographic stack $\underline{\boldsymbol{v}}$ can be modeled by the following (complex-valued) linear model:

$$\underline{\boldsymbol{v}} = \underline{\boldsymbol{\Phi}}\,\underline{\boldsymbol{u}} + \underline{\boldsymbol{\epsilon}} \tag{8}$$

where $\underline{\boldsymbol{\epsilon}}$ stands for the noise. This corresponds to a generalization of the tomographic direct model (3) where we additionally consider the geometric transformation from ground geometry to SAR geometry and model at once the measurements for all the pixels. In order to invert this tomographic model, it is necessary to introduce some regularization terms.

B. Spatial regulariations for tomographic inversion

It is often desired to reconstruct volumes with a discretization in heights that is finer than the resolution given by the synthetic aperture in the height direction (i.e., super-resolution). The inversion of equation (8) is therefore ill-posed (more unknowns than measurements) and requires some regularization. In the following, we denote $\mathcal{R}$ the regularization function. Since the intrinsic phase of a scatterer is typically modeled as uniformly distributed and independent from one scatterer to another, no specific regularization can be enforced on the phase of our unknown complex reflectivities $\underline{\boldsymbol{u}}$. We define the regularization $\mathcal{R}$ as a function of the modulus of $\underline{\boldsymbol{u}}$ only. The reconstruction of the volume of complex reflectivities $\hat{\underline{\boldsymbol{u}}}$ is thus obtained by solving an optimization problem of the general form:

$$\hat{\underline{\boldsymbol{u}}} = \arg\min_{\underline{\boldsymbol{u}}} \frac{1}{2}||\underline{\boldsymbol{\Phi}}\,\underline{\boldsymbol{u}} - \underline{\boldsymbol{v}}||_2^2 + \mathcal{R}(|\underline{\boldsymbol{u}}|) \tag{9}$$

To design the regularization function $\mathcal{R}$, we need to select a function that favors volumes of reflectivities $|\underline{\boldsymbol{u}}|$ that are often present in urban environments. Many different such functions could be considered, we selected a function based on the two following remarks:

- The 3-D scene can be represented as a sparse volume. Indeed a good reconstruction should retrieve only the illuminated part of the buildings and of the ground. The estimated volume is then mostly filled with zero intensity voxels.
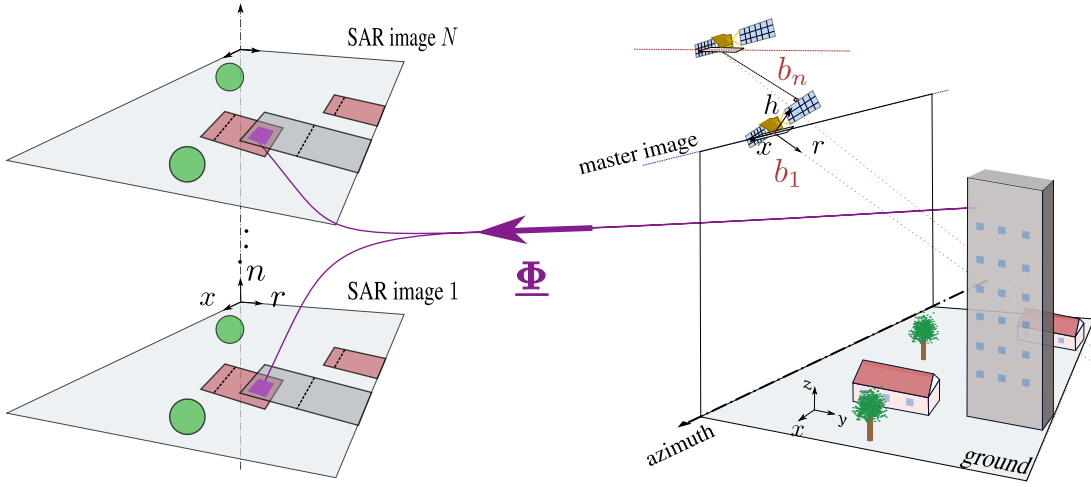
Fig. 2. The operator $\underline{\boldsymbol{\Phi}}$ performs a projection of voxels from the 3-D scene in ground geometry on the SAR tomographic stack.

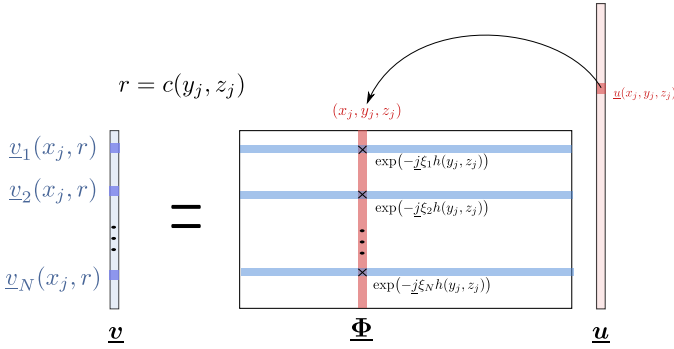Fig. 3. Construction of the matrix $\underline{\boldsymbol{\Phi}}$. The radar cell associated to each position of the 3-D scene is computed as well as the phase term corresponding to distance between a point in 3-D space and an antenna.

- The illuminated structures are spatially smooth (continuous surfaces: frontages, rooftops, ground).

These remarks suggest the following regularization function:

$$\forall \boldsymbol{w} \in \mathbb{R}^{N_x.N_y.N_z}, \ \mathcal{R}(\boldsymbol{w}) = \frac{\mu_{\text{x}}}{2}\|\boldsymbol{D}_x\boldsymbol{w}\|_2^2 + \frac{\mu_{\text{y}}}{2}\|\boldsymbol{D}_y\boldsymbol{w}\|_2^2$$
$$+ \frac{\mu_{\text{z}}}{2}\|\boldsymbol{D}_z\boldsymbol{w}\|_2^2 + \mu_{\ell_1}\|\boldsymbol{w}\|_1 \quad (10)$$

where the matrices $\boldsymbol{D}_x$, $\boldsymbol{D}_y$ and $\boldsymbol{D}_z$ stand for the finite differences operators in the x, y and z directions, and parameters μ_x, μ_y, μ_z and μ_{ℓ_1} weight each term. The ℓ_1 norm favors vectors $\underline{\boldsymbol{w}}$ with many zeros while the terms with the finite difference operators enforce a spatial smoothness. As μ_{ℓ_1} controls this behavior, it plays a crucial role in the reconstruction.

In urban environment, dihedral and trihedral structures produce very strong echoes. When only an ℓ_1 norm is minimized, it is hard to reconstruct at the same time very strong scatterers and weaker scatterers on the ground or rooftops. The role of the first three terms is thus to favor spatial smoothness and hence preserve these scatterers whenever they are close to other scatterers. In urban areas most of the buildings are supposed to show straight walls and then a natural direction of elongation along the z axis. Of course the true orientation of the objects is a priori unknown and it is most likely that

they may not always follow the x and y directions. However, the conjoint minimization of the ℓ_1 norm of the reflectivity and the square ℓ_2 norm of its gradient tends to favor sparse and continuous reflectivity profile. The obtained reconstructions present then less outliers to the benefit of structural elements.

When reconstructing images of urban areas, we found that it was impossible to find a value of the sparsity parameter μ_{ℓ_1} that would both preserve the weakest scatterers and successfully suppress side-lobes in areas with strong scatterers. To improve the reconstructions, we introduced a spatially variant regularization based on the square-root of the estimated intensity of the master image which can be obtained using a denoising algorithm such as NL-SAR [7] or simply using the average intensity depending of the configuration of the acquisitions. We define the diagonal matrix $\boldsymbol{D}_{\ell_1}$ whose j-th diagonal entry is equal to the square-root of the estimated intensity at the corresponding azimuth and range coordinates (i.e., such that the range r verifies $r - \frac{\delta_r}{2} < \rho_{1;y_j,z_j} < r + \frac{\delta_r}{2}$). The equation (10) is then modified into:

$$\forall \boldsymbol{w} \in \mathbb{R}^{N_x.N_y.N_z}, \ \mathcal{R}(\boldsymbol{w}) = \frac{\mu_{\text{x}}}{2}\|\boldsymbol{D}_x\boldsymbol{w}\|_2^2 + \frac{\mu_{\text{y}}}{2}\|\boldsymbol{D}_y\boldsymbol{w}\|_2^2$$
$$+ \frac{\mu_{\text{z}}}{2}\|\boldsymbol{D}_z\boldsymbol{w}\|_2^2 + \mu_{\ell_1}\mathbf{1}^T\boldsymbol{D}_{\ell_1}\boldsymbol{w} \quad (11)$$

where $\mathbf{1}$ is the vector of size $N_x.N_y.N_z$ with each entry equal to 1.

C. Optimization algorithm for the 3-D reconstruction

The minimization problem (9) is not easy. It is indeed large scale (millions up to several billions unknowns) and non-convex. The non-convexity is illustrated in a simple case where the vector $\underline{\boldsymbol{u}}$ has only two elements in Fig. 4: the spatial smoothness favors vectors such that the modulus of each entry is close. Since the regularization is independent on the phase, the minimum (for a fixed value of $\underline{u}_1$) corresponds to a set of complex values with equal modulus (the white circle drawn on Fig. 4).

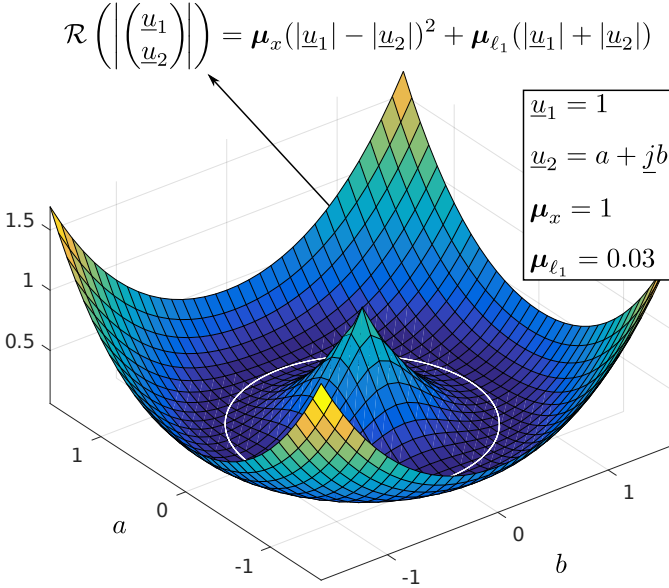To solve the minimization problem (9), we use a variable splitting approach in order to break down the problem into a

$$\mathcal{R}\left(\left|\begin{pmatrix}\underline{u}_1\\\underline{u}_2\end{pmatrix}\right|\right) = \boldsymbol{\mu}_x(|\underline{u}_1| - |\underline{u}_2|)^2 + \boldsymbol{\mu}_{\ell_1}(|\underline{u}_1| + |\underline{u}_2|)$$

$\underline{u}_1 = 1$

$\underline{u}_2 = a + jb$

$\boldsymbol{\mu}_x = 1$

$\boldsymbol{\mu}_{\ell_1} = 0.03$

Fig. 4.   Illustration of the non-convexity of spatial regularizations expressed on the modulus of the complex reflectivities.

sequence of simpler problems. We introduce two new vectors: $\underline{f}$ (complex-valued: $\underline{f} \in \mathbb{C}^{(N_x \cdot N_y \cdot N_z)}$) and $\boldsymbol{w}$ (real-valued: $\boldsymbol{w} \in \mathbb{R}^{+(N_x \cdot N_y \cdot N_z)}$). Problem (9) is formally equivalent to the following constrained problem:

$$\hat{\underline{u}} = \arg\min_{\underline{u}} \frac{1}{2}\|\boldsymbol{\Phi}\,\underline{u} - \underline{v}\|_2^2 + \mathcal{R}(\boldsymbol{w}) \qquad (12)$$

$$\text{s.t.} \begin{cases} \underline{u} = \underline{f} \\ |\underline{f}| = \boldsymbol{w} \end{cases}$$

To solve this constrained optimization problem, we apply the method of multipliers [30], i.e., we look for a saddle point of the augmented Lagrangian $\mathcal{L}$ defined by:

$$\mathcal{L}(\underline{u}, \underline{f}, \boldsymbol{w}, \underline{d}_1, d_2) = \frac{1}{2}\|\boldsymbol{\Phi}\underline{u} - \underline{v}\|_2^2 + \frac{\beta_1}{2}\|\underline{f} - \underline{u} + \underline{d}_1\|_2^2$$
$$+ \frac{\beta_2}{2}\left\|\boldsymbol{w} - |\underline{f}| + d_2\right\|_2^2 + \mathcal{R}(\boldsymbol{w}) \qquad (13)$$

where $\underline{d}_1 \in \mathbb{C}^{(N_x \cdot N_y \cdot N_z)}$ and $d_2 \in \mathbb{R}^{(N_x \cdot N_y \cdot N_z)}$ are the scaled dual variables and β_1 and β_2 are penalty parameters (relevant only to the optimization method, i.e., impacting the convergence). We follow the hierarchical approach described in [31] and minimize jointly on the variables $\underline{u}$ and $\boldsymbol{w}$ while $\underline{f}$ is substituted with its optimal value $\underline{f}^*(\underline{u}, \boldsymbol{w})$:

$$\underline{f}^* = \arg\min_{\underline{f}} \frac{\beta_1}{2}\|\underline{f} - \underline{u} + \underline{d}_1\|_2^2 + \frac{\beta_2}{2}\left\|\boldsymbol{w} - |\underline{f}| + d_2\right\|_2^2$$
$$\qquad (14)$$
$$= \left[\frac{\beta_1 \cdot |\underline{u} - \underline{d}_1| + \beta_2 \cdot (\boldsymbol{w} + d_2)}{\beta_1 + \beta_2}\right]^+ \exp\left[\mathbf{j} \cdot \arg(\underline{u} - \underline{d}_1)\right]$$
$$\qquad (15)$$

where $.^+$ is the identity on $[0, +\infty[$ and is the constant null function on $]-\infty, 0[$.

Proof. The second term in (14) only depends on the modulus of $\underline{f}$. The phase of $\underline{f}^*$ is then driven by the first term and

must be chosen equal to that of $\underline{u} - \underline{d}_1$ so as to minimize the cost function. There remains to estimate the modulus of $\underline{f}^*$ which is a solution of a 1-D quadratic problem

$$\arg\min_{\rho \geq 0} \frac{\beta_1}{2}\||\boldsymbol{\rho} - |\underline{u} + \underline{d}_1|\|_2^2 + \frac{\beta_2}{2}\|\boldsymbol{w} - \boldsymbol{\rho} + d_2\|_2^2 \quad (16)$$

ρ is either given by the unconstrained solution:

$$\rho^* = (\beta_1 \cdot |\underline{u} - \underline{d}_1| + \beta_2 \cdot (\boldsymbol{w} + d_2))/(\beta_1 + \beta_2),$$

or $\rho = 0$. The optimal value of ρ and of the phase together lead to the expression (15). $\qquad \square$

This leads to the following algorithm[1] which alternates between the joint minimization with respect to variables $\underline{u}$ and $\boldsymbol{w}$, and the update of dual variables.

Algorithm Tomographic SAR 3-D Inversion

Input: $\underline{v}$ (stack of SLC SAR images)
Output: $\hat{\underline{u}}$ (3-D cube of complex reflectivities)
   Initialization :
1: $\hat{\underline{d}}_1 \leftarrow \mathbf{0}$
2: $\hat{d}_2 \leftarrow \mathbf{0}$
3: **while** not converged **do**
4:    $\{\hat{\underline{u}}, \hat{\boldsymbol{w}}\} \leftarrow \texttt{approximate\_min}(\underline{v}, \hat{\underline{d}}_1, \hat{d}_2, \hat{\underline{u}}, \hat{\boldsymbol{w}})$
5:    $\hat{d}_2 \leftarrow \hat{d}_2 + \hat{\boldsymbol{w}} - |\underline{f}^*(\hat{\underline{u}}, \hat{\boldsymbol{w}})|$
6:    $\hat{\underline{d}}_1 \leftarrow \hat{\underline{d}}_1 + \underline{f}^*(\hat{\underline{u}}, \hat{\boldsymbol{w}}) - \hat{\underline{u}}$
7: **end while**
8: **return** $\hat{\underline{u}}$

Procedure $\texttt{approximate\_min}$
Input: $\underline{v}$ (stack of SLC SAR images)
Input: $\hat{\underline{d}}_1$ (vector of dual variables)
Input: $\hat{d}_2$ (vector of dual variables)
Input: $\hat{\underline{u}}$ (current vector of primal variables)
Input: $\hat{\boldsymbol{w}}$ (current vector of primal variables)
Output: $\{\hat{\underline{u}}, \hat{\boldsymbol{w}}\}$ (approximate solution)

1: $\mathscr{C}(\underline{v}, \hat{\underline{d}}_1, \hat{d}_2, \hat{\underline{u}}, \hat{\boldsymbol{w}}) = \frac{1}{2}\|\boldsymbol{\Phi}\underline{u} - \underline{v}\|_2^2 + \mathcal{R}(\boldsymbol{w})$
    $+ \frac{\beta_1}{2}\|\underline{f}^*(\underline{u}, \boldsymbol{w}) - \underline{u} + \hat{\underline{d}}_1\|_2^2 + \frac{\beta_2}{2}\left\|\boldsymbol{w} - |\underline{f}^*(\underline{u}, \boldsymbol{w})| + \hat{d}_2\right\|_2^2,$

2: $\nabla_{\underline{u}}\mathscr{C} = \boldsymbol{\Phi}^H(\boldsymbol{\Phi}\underline{u} - \underline{v}) + \beta_1(\underline{u} - \underline{f}^*(\underline{u}, \boldsymbol{w}) - \hat{\underline{d}}_1)$

3: $\nabla_{\boldsymbol{w}}\mathscr{C} = (\mu_x \boldsymbol{D}_x^T \boldsymbol{D}_x + \mu_y \boldsymbol{D}_y^T \boldsymbol{D}_y + \mu_z \boldsymbol{D}_z^T \boldsymbol{D}_z)\boldsymbol{w} + \mu_{\ell_1}\mathbf{1}$
           $+ \beta_2(\boldsymbol{w} - \underline{f}^*(\underline{u}, \boldsymbol{w}) + \hat{d}_2).$

4: $\nabla\mathscr{C}(\underline{v}, \hat{\underline{d}}_1, \hat{d}_2, \hat{\underline{u}}, \hat{\boldsymbol{w}}) = \begin{pmatrix}\nabla_{\underline{u}}\mathscr{C}\\\nabla_{\boldsymbol{w}}\mathscr{C}\end{pmatrix},$

    Call quasi-Newton minimization algorithm :
5: $\{\hat{\underline{u}}, \hat{\boldsymbol{w}}\} \leftarrow \texttt{quasi\_Newton}[\mathscr{C}(\underline{v}, \hat{\underline{d}}_1, \hat{d}_2, \hat{\underline{u}}, \hat{\boldsymbol{w}}),$
          $\nabla\mathscr{C}(\underline{v}, \hat{\underline{d}}_1, \hat{d}_2, \hat{\underline{u}}, \hat{\boldsymbol{w}}),$
          $\texttt{constraint: } \hat{\boldsymbol{w}} \geq 0]$

[1]compared to the well-known alternating directions method of multipliers (ADMM) [32], this hierarchical approach has been shown in [31] to converge faster and to be less sensitive to the tuning of the optimization parameters β_1 and β_2; moreover, we improve the convergence by constraining $\boldsymbol{w}$ to be positive, such a constraint would be costly to enforce with ADMM.
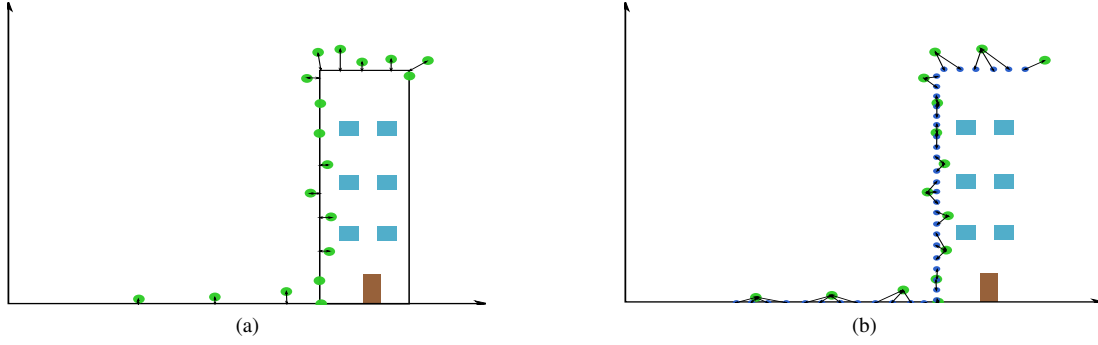
Fig. 5. Accuracy (a) et completeness (b) criteria. The accuracy computes the mean of the error for each estimated point. The completeness gives an indication on the proportion of holes in the rendering.

After replacing $\boldsymbol{f}$ by its optimal value $\boldsymbol{f}^*(\underline{\boldsymbol{u}}, \boldsymbol{w})$ and constraining $\boldsymbol{w}$ to be positive, the objective function is differentiable (since $\boldsymbol{w} \geq \boldsymbol{0}$, $\|\boldsymbol{w}\|_1 = \boldsymbol{1}^T \boldsymbol{w}$, which is differentiable; moreover, after substituting $\boldsymbol{f}$ by the optimal value $\boldsymbol{f}^*$ in the quadratic terms, the cost function becomes smooth, see [31]). The cost function can thus be minimized using a limited memory quasi-Newton algorithm that handles positivity constraints, such as L-BFGS-B [33], or the slightly more efficient algorithm VMLM-B [34] that we used, with Eric Thiébaut's freely available implementation[2]. The minimization step described in Procedure `approximate_min` does not need to be performed up to a high precision. A few (e.g., ten) iterations of the quasi-Newton algorithm are sufficient since the algorithm is warm-restarted. In the definition of the gradient of the cost function (lines 2 to 4 of Procedure `approximate_min`), we used the fact that $\frac{\partial \mathscr{C}}{\partial [\boldsymbol{f}^*]_i} \frac{\partial [\boldsymbol{f}^*]_i}{\partial [\boldsymbol{u}]_j} = \frac{\partial \mathscr{C}}{\partial [\boldsymbol{f}^*]_i} \frac{\partial [\boldsymbol{f}^*]_i}{\partial [\boldsymbol{w}]_k} = 0$ for all i, j, k, even though $\boldsymbol{f}^*$ depends on variables $\underline{\boldsymbol{u}}$ and $\boldsymbol{w}$ when applying the chain rule since $\frac{\partial \mathscr{C}}{\partial [\boldsymbol{f}^*]_i} = 0$ for all i. The necessary condition for $\boldsymbol{f}^*$ being a minimizer of the augmented Lagrangian is then respected. As the function we want to minimize is non-convex, the convergence is assumed when we met a local minimum. In practice the algorithm can then be stopped after a fixed number of iterations (around 60) or when the distance between two iterations is below a given threshold.

The priors proposed in this section are designed to be simple and to lead to a function $\mathcal{R}$ that is easy to optimize. In the following experiments, the relevance of this method is shown by comparing to other state of the art estimators on simulated and real data. The proposed framework is very general and can easily be adapted to include other spatial regularizations expressed in ground coordinates.

IV. EXPERIMENTS

A. Evaluation Protocol

In this part we compare 3-D scenes reconstructed by five different approaches: classical beamforming [4], Capon beamforming [22], MUSIC [24], CS [12] [11] and our algorithm. Evaluation of different SAR tomographic estimators can be hard as they originate from various fields and are thus not

[2]https://github.com/emmt/OptimPackLegacy

suited to the same applications. Moreover, they do not always estimate the same quantities: classical and Capon beamforming give the estimated power of the scatterers, CS and our 3-D inversion estimate the scatterers reflectivity, and MUSIC only provides us with their position. Here, we will only evaluate the quality of the localization, as all methods can be used to estimate the 3D location of scatterers and this represents a rich information in urban areas. It can be compared to a ground truth (digital elevation model).

We use the accuracy and completeness criteria introduced in [19] to evaluate the tomographic SAR reconstructions. Those two metrics are defined for the evaluation of point clouds reconstructions. Tomographic inversions must thus be converted into a discrete representation in order to compute these performance criteria.

1) Accuracy: For a given discrete reconstruction $\hat{\mathcal{P}}$, the accuracy represents the mean distance from each point in $\hat{\mathcal{P}}$ to the ground truth $\mathcal{P}$.

$$A(\hat{\mathcal{P}}, \mathcal{P}) = \frac{1}{N_{\hat{p}}} \sum_{j=1}^{N_{\hat{p}}} \min_k \|\hat{p}_j - p_k\|_2 \qquad (17)$$

where $\hat{p}_j \in \hat{\mathcal{P}}$ is the jth point of the estimated point cloud $\hat{\mathcal{P}}$ and $p_k \in \mathcal{P}$ is the kth point of the ground truth. $N_{\hat{p}}$ is the number of points in the estimated reconstruction. Accuracy indicates whether reconstructed points are correctly located.

2) Completeness: The completeness corresponds to the mean distance from each point of the ground truth to the points in $\hat{\mathcal{P}}$:

$$C(\hat{\mathcal{P}}, \mathcal{P}) = \frac{1}{N_p} \sum_{k=1}^{N_p} \min_j \|\hat{p}_j - p_k\|_2, \qquad (18)$$

with N_p the number of points in the ground truth. Completeness indicates whether the ground truth is well represented by the set of points in the reconstruction.

These two metrics are illustrated in Fig. 5 where green dots represent estimated points $\hat{p}_j$ and blue squares the ground truth points p_k. Accuracy and completeness provide complementary information: accuracy improves when reconstructed points are close to actual points from the ground truth, but does not

indicate when points are missing (holes in the reconstruction). For instance, if we could retrieve a single point of the scene (the strongest permanent scatterer, for example) with a location very close to the ground truth, the accuracy would be excellent ($A(\hat{\mathcal{P}}, \mathcal{P})$ near zero) while the completeness $C(\hat{\mathcal{P}}, \mathcal{P})$ would be large, indicating that much of the scene is missing in the reconstruction. Conversely, if we loosely select points (many points for each resolution cell), we would obtain a dense volume, thus a good completeness (low completeness value $C(\hat{\mathcal{P}}, \mathcal{P})$), but erroneously selected points lying far from the true surfaces would lead to a poor accuracy score (large accuracy value $A(\hat{\mathcal{P}}, \mathcal{P})$).

The accuracy and completeness can be jointly used to evaluate the quality of a discrete reconstruction. Beforehand, 3-D points must be extracted from the volume of voxels obtained by the tomographic reconstruction algorithms. We performed this extraction by finding the local maxima in each radar resolution cell. This produces a first point-based representation of the data. Then, to reduce the sensitivity to noise and side-lobes, we applied a threshold on the obtained point cloud. Starting from a reconstructed volume $\hat{\underline{u}}$, we obtain the collection of 3-D points $\hat{\mathcal{P}}_{\hat{\underline{u}},t}$:

$$\hat{\mathcal{P}}_{\hat{\underline{u}},t} = \left\{ p_j = (x_j, y_j, z_j) \in \mathbb{R}^3, \text{ such that} \right.$$

$$\left. |\hat{\underline{u}}_{x_j,y_j,z_j}| \geq |\hat{\underline{u}}_{x_j,\nu_y,\nu_z}| \text{ and } |\hat{\underline{u}}_{x_j,y_j,z_j}| > t \right\} \quad (19)$$

where ν_y and ν_z are the coordinates of the direct neighboring voxels restricted to the same radar resolution cell and t is a threshold. In the following experimentss, when we visually compare reconstructions obtained by different algorithms, we always select the threshold t as the value t^* that achieves the best trade-off between accuracy and completeness:

$$t^* = \arg\min_t A(\hat{\mathcal{P}}_{\hat{\underline{u}},t}, \mathcal{P})^2 + C(\hat{\mathcal{P}}_{\hat{\underline{u}},t}, \mathcal{P})^2. \quad (20)$$

When evaluating different tomographic estimators, we are generally interested in different regimes (accurate reconstruction of the strongest scatterers, dense reconstruction of most scatterers). To capture these different cases, rather than considering a single reconstruction obtained with threshold t^*, we also represent the accuracy score $A(\hat{\mathcal{P}}_{\hat{\underline{u}},t}, \mathcal{P})$ as a function of the completeness $C(\hat{\mathcal{P}}_{\hat{\underline{u}},t}, \mathcal{P})$. Applications that focus on the reconstruction of permanent scatterers will favor algorithms that achieve the highest accuracy values (even if the completeness is poor). Many works have considered linear dependencies of the phase of permanent scatterers on various parameters such as their height but also their speed or the temperature [35] [28]. If the reconstruction of surfaces is the aim, completeness should be favored, even if this degrades the accuracy.

B. Simulated Data

In this section, we compare our 3-D tomographic reconstruction algorithm to state-of-the-art methods, using simulated data. We simulate two stack of images corresponding to two different configurations of trajectories: a regular sampling along the elevation axis h (referred to in the following as Reg), and the actual trajectories of TerraSAR-X given in Fig. 17 (referred to in the following as TSX). As in [12], the decorrelation effect is introduced by adding a Gaussian white noise and the phase noise is modeled by adding a uniformly distributed random phase between $-\pi$ and π. Fig. 6 shows the theoretical distribution of the scatterers. Both scenes are composed of a ground at a constant altitude, a wall and a roof, resulting in a large layover area in the SAR images. The simulated building is higher than the estimated elevation resolution (as given by Fourier inversion) in the TSX case. The reflectivity of the scatterers is set constant on all the scene. The additive noise level is of -0.6 dB and is the same for both experiments. In the Reg case, the scatterers are well separated in the azimuth and range directions. In the TSX experiment, the density of scatterers is bigger for each cell resulting in clusters of neighboring scatterers being projected in the same radar cell. The signal is thus corrupted with a speckle effect induced by the coherent addition of the back-scatterered signals. Due to the relatively strong additive noise and the geometrical decorelation the average coherence is 0.63 for the Reg images and 0.68 for the TSX ones. Finally as the resolution is not the same in the two experiments the TSX structure is taller and the resulting images have a bigger size in range.

C. Influence of the regularization parameters

To show the influence of each regularization parameter, we performed different reconstructions of the Reg scene for various sets of regularization values (in particular, with one parameter chosen so as to illustrate the effect of either under-regularization or over-regularization). In these expreriments, the phase noise is not present to show clearly the structural influence of the spatial smoothing. The results are presented in figures 7, 8, 9 and 10. On this well-sampled scene with a good resolution, the parameter with the largest influence is μ_{ℓ_1}, associated to the sparsity constraint. Fig. 7 illustrates that when μ_{ℓ_1} is too large, there are some holes in the reconstruction, while a value of μ_{ℓ_1} that is too small leaves side-lobes and outliers. The effect of over-smoothing (values of μ_x, μ_y or μ_z too large) is the extension of structures in the direction of the spatial smoothing. This is visible in particular with structures whose orientation differs from that of the smoothing, see in Fig. 9 the widening of the wall due to excessive smoothing in the horizontal direction. The same effect appears on the rooftop in Fig. 10. Insufficient smoothing translates into residual fluctuations (i.e., large variance) that are reduced by increasing spatial smoothing, see in particular the corner between the roof and the wall, or the ground and the wall in Fig. 9 and 10. Outliers located far from the actual surfaces can also be observed when the spatial regularization is too weak.

By combining sparsity and spatial smoothness constraints, our algorithm is very flexible and well-suited both to the reconstruction of urban areas where scatterers are sparsely distributed over 2-D surfaces and to vegetated areas where scatterers are spread in volumes in a smooth 3-D distribution. The downside of this flexibility is the necessity to tune
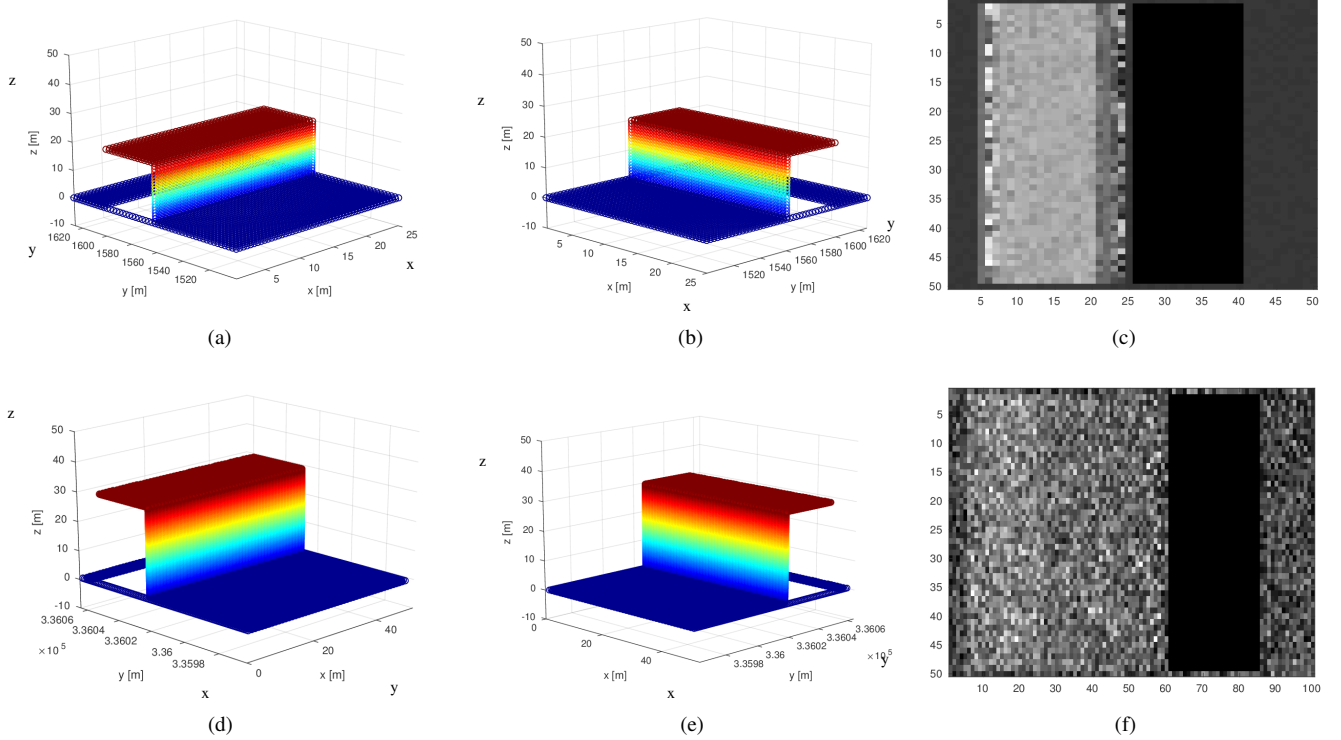
(a)

(b)

(c)

(d)

(e)

(f)

Fig. 6. Two simulations of a simple 3-D scene: first row, considering a regular sampling of trajectories along the h axis; second row: using the trajectories of TerraSAR-X satellites. Sub-figures (a), (b), (d) and (e) show two views of the scenes while sub-figures (c) and (f) correspond to the average of the simulated SAR images. The sub-figures in the top row correspond to the `Reg` case and the ones in the bottom correspond to the `TSX` case. The `Reg` case consists of 20 images acquired with uniformly distributed sensors along the elevation axis. In the `TSX` case the geometry of the scene is given by the true TerraSAR-X satellites position for each of the 40 acquisitions. In both case the SNR is 1.7dB. In the `TSX` experiment, more scatterers are projected in the same resolution cell which produces a important speckle effect.

four regularization parameters. In numerical simulations, the ground-truth can be used to select the set of regularization parameters $\{\mu_x, \mu_y, \mu_z, \mu_{\ell_1}\}$ that offers the best performance, as measured by the Minimum Accuracy / Completeness Trade-off (MACT) $\min A(\hat{\mathcal{P}}_{\hat{\boldsymbol{u}}, t^*}, \mathcal{P})^2 + C(\hat{\mathcal{P}}_{\hat{\boldsymbol{u}}, t^*}, \mathcal{P})^2$. The regularization parameters can then be tuned in order to reach the best possible trade-off. We advice to perform the one dimensional optimization, by importance: μ_{ℓ_1}, then μ_z, then μ_x and μ_y. While tuning the sparsity parameter, the other smoothing parameters should be put to 0.

Other derivative free methods can be used to fix all the parameters at once. We compared the described alternating minimization strategy with Nelder-Mead's simplex method by initializing with different sets of parameters and found similar values for the lowest error. The differences for μ_{ℓ_1} is of 21% and below 14% for the smoothing parameters. As being more precise, this approach is highly computational and needs to be reinitialize several times due to the non-convexity of the error with respect to the parameters. In the absence of ground truth, a simple numerical simulation using the same geometrical configuration and sensor noise power can be generated in order to automatically tune the parameters. If a simulated scene isn't an option, the alternating minimization strategy can still be applied in an interactive fashion: the user tunes each parameter, by order of importance, in order to reach a satisfying reconstruction.

To further illustrate the behavior of our algorithm with respect to its regularization parameters, we plotted accuracy as a function of completeness for different sets of parameters. These curves are drawn for our two simulation cases `Reg` and `TSX` respectively in Fig. 11 and 12. We observe that the best accuracy /completeness trade-off (point of the curve closest to the origin of the axes) is reached for a unique set of parameters that can thus be found for example by binary search.

The optimization parameters β_1 and β_2 have an impact on the convergence speed. We found that, when starting from a volume initialized at zero, using large penalty parameters β_1 and β_2 produces very quickly a sparse reconstruction while lower contrasted structures are correctly reconstructed after many more iterations. Smaller values of the parameters help to reconstruct those structures, at the cost of a slower convergence (i.e., sidelobes suppression) in the brightest areas. Penalty parameters β_1 and β_2 can be set according to methods described in [32], [36]. Here, we fixed those parameters to 10 after having tested different values.

D. Comparaison of different tomographic estimators

The figures 13 and 14 give the 3-D reconstructions obtained using several state-of-the-art tomographic reconstruction methods and our spatially regularized approach: (a) classical beamforming, (b) Capon beamforming, (c) the parametric estimators MUSIC, (d) Compressed Sensing, and (e) our
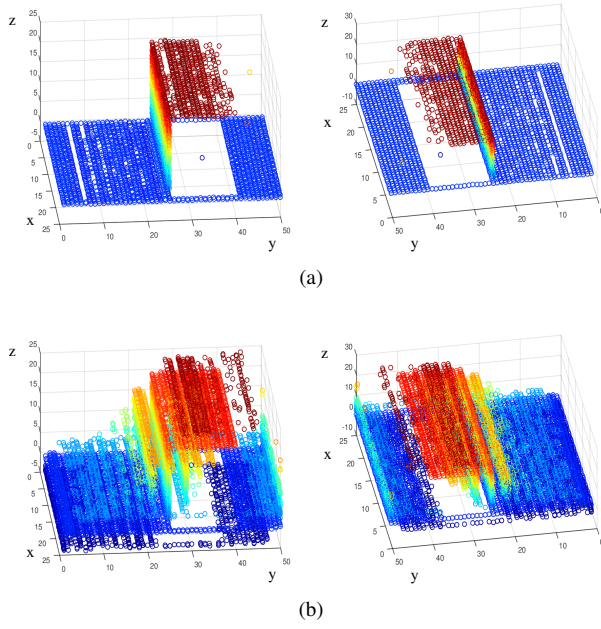
(a)

(b)

Fig. 7. Influence of the parsimony μ_{ℓ_1} parameter. In (a) the parameter is way above its optimal value resulting in a lot of holes in the structure. In (b) we use a small value of μ_{ℓ_1} resulting in a high number of outliers.
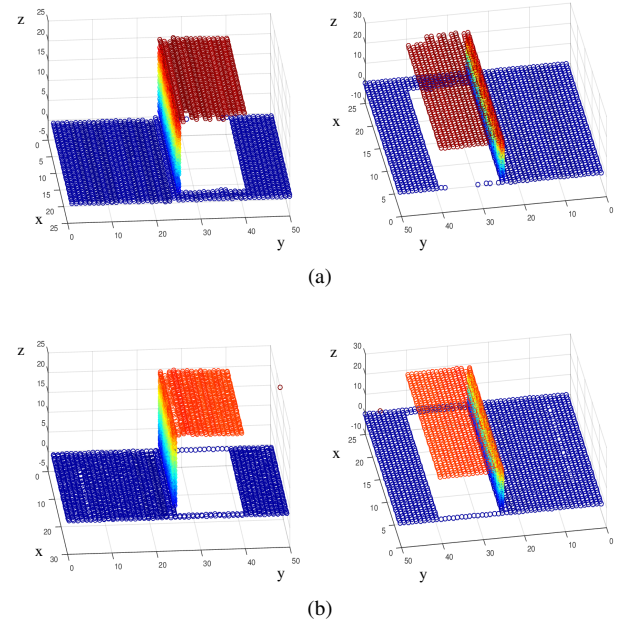

(a)

(b)

Fig. 8. Influence of the μ_x parameter. In (a) the parameter is way above its optimal value. In (b) the value of the parameter is set to zero.
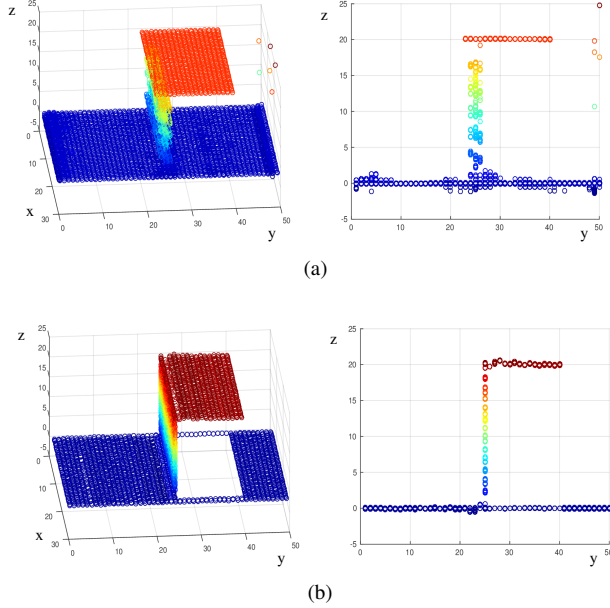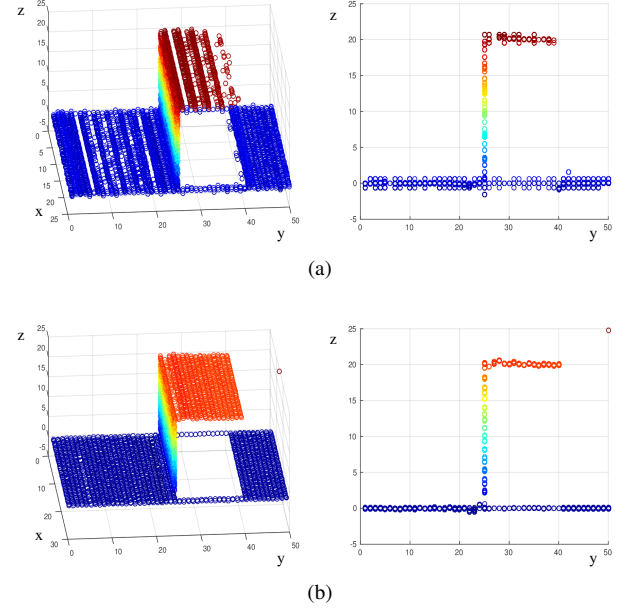

(a)

(b)

Fig. 9. Influence of the μ_y parameter. In (a) the parameter is way above its optimal value. In (b) the value of the parameter is set to zero.

(a)

(b)

Fig. 10. Influence of the μ_z parameter. In (a) the parameter is way above its optimal value. In (b) the value of the parameter is set to zero.

algorithm. The corresponding curves of accuracy as a function of completeness are presented in figure 15 for the two scenes.

In the easy case where the antennas are uniformly distributed classical beamforming presents good performances with respect to the accuracy-completeness metrics, achieving a smaller error than Capon beamforming or MUSIC. In the TSX case where the theoretical resolution is worst, few of the wall is retrieved leading to a poor completeness score. A closer inspection of the TSX reconstructions indicates that the presence of side-lobes (*i.e.* many outliers).

Both Capon beamforming and MUSIC, which rely on the covariance matrix to achieve super resolution, manage to sup-

press most of the outliers. However, the averaging introduced for the covariance matrix estimation produces an extension of the ground and walls. For the Reg case, this explains the bounded accuracy scores even when the completeness is poor. Capon beamforming manages to reduce the side lobes in the Reg case but not in the TSX simulation where the irregular sampling produces dramatic biasing effects on the estimation of the position of the scatterers.

In the TSX scene, the best accuracy / completeness trade-off from the MUSIC pseudo-spectra produces a sparse scene with no outliers. However the bias present in the estimation of the covariance matrix strongly deforms the wall and introduces
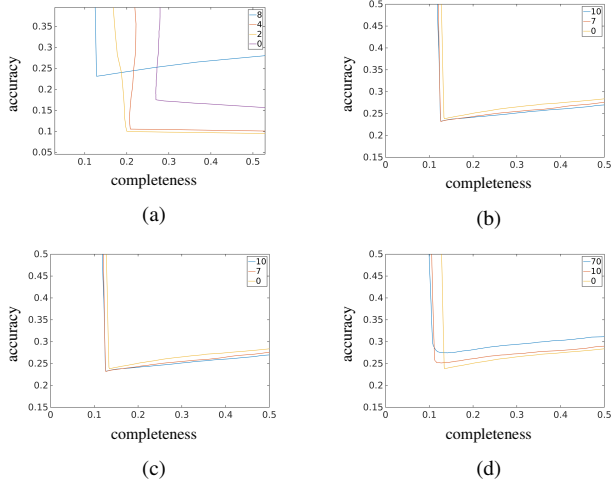
Fig. 11. Study of the influence of each parameter on the metrics for the `Reg` scene (cf. Fig. 6 (a)). The plots in (a), (b), (c) and (d) correspond respectively to different values of μ_{ℓ_1}, μ_x, μ_y and μ_z with the three other parameters being fixed.
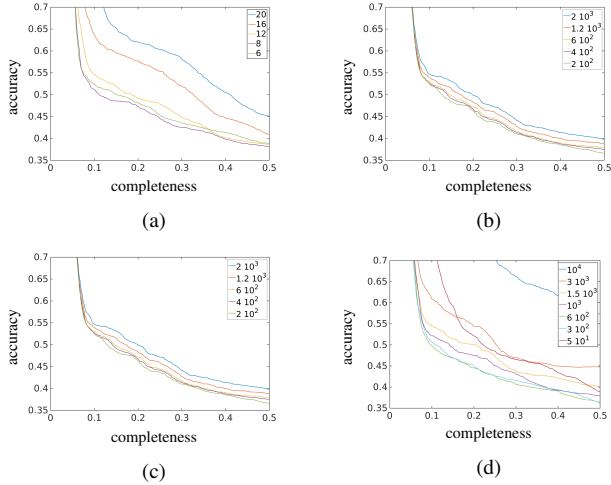

Fig. 12. Study of the influence of each parameter on the metrics for the `TSX` scene (cf. Fig. 6 (a)). The plots in (a), (b), (c) and (d) correspond respectively to different values of μ_{ℓ_1}, μ_x, μ_y and μ_z with the three other parameters being fixed.

some fake discontinuities.

All the spectral estimators presented (classical beamforming, Capon beamforming and MUSIC) suffer from the averaging step needed to estimate the covariance matrix, which produces visibly distorted walls and/or rooftop in the `TSX` reconstruction. An elongation of the rooftop and a reduction of the ground is also visible in the `Reg` case. However the naive covariance estimation step makes those estimators efficient for uniform areas which is illustrated by the generally accurate ground reconstruction.

As expected, CS is one of the top estimator in accuracy and often manages to achieve the most accurate representations. However, the precision comes with the cost of a loss of information which explains the relatively bad scores in completeness.

The proposed 3-D inversion algorithm shows improved performances compared to the CS approach. Some outliers that

were present even when imposing an ℓ_1 penalty are suppressed thanks to the spatial smoothness prior. As the method uses a sparse prior, some information may be lost but thanks to the geometrical constraints, most of the rooftop and ground are filled. Moreover, in the `TSX` case the wall and rooftop are better located around their true position and the reconstruction is more robust to deformation introduced by the proximity of the scatterers.

The tuning of the parameter in the simulated experiences has been done by exploiting the knowledge of the scene. Of course, when dealing with real data, this information is not available. Then, to test if the proposed method is able to exhibit better performances when the tuning of the parameter is not ideal, the model parameter are modified with increasing perturbation values. The results are shown in the table I. Up to an isotropic perturbation of 50, the proposed method still shows a better Minimum Accuracy-Completeness Tradeoff (MACT) than the other tested methods. An other concern is whether the parameters need to be adapted to the scatterers power. Real-data experiments conducted over a very heterogeneous scene indicates that the reconstruction doesn't present artifacts due to excessive smoothing in some areas. To confirm this observation with a controlled experiment, reconstruction scores of the `TSX` scene with random scatterers amplitude are also tested. The amplitudes were chosen randomly for each azimuth in a given range. The different experiments correspond increasing ranges. The results of the experiments are shown in table II. The proposed method still performs better, even when the amplitudes of the scatterers varies by 3 orders of magnitude and the regularization parameters are kept constant.

Estimator	MACT
Beamforming	0.96
Capon Beamforming	0.98
MUSIC	0.66
CS	0.71
3-D inversion ($\boldsymbol{\mu}^*$)	**0.57**
3-D inversion ($\boldsymbol{\mu}^* - 50\%$)	0.62
3-D inversion ($\boldsymbol{\mu}^* - 100\%$)	0.62
3-D inversion ($\boldsymbol{\mu}^* + 50\%$)	0.59
3-D inversion ($\boldsymbol{\mu}^* + 100\%$)	0.65
3-D inversion ($\boldsymbol{\mu}^* + 200\%$)	0.74

TABLE I

EVOLUTION OF THE MINIMUM ACCURACY-COMPLETENESS TRADEOFF (MACT) WHEN THE PARAMETERS ARE GETTING FAR FROM THEIR OPTIMAL VALUE $\boldsymbol{\mu}^* = \{\mu_{\ell_1}^*, \mu_x^*, \mu_y^*, \mu_z^*\}$. THE SMALLER THE MACT THE BETTER.

Estimator/Amplitudes	[1, 1]	[10^{-1}, 10]	[10^{-2}, 10^2]	[10^{-3}, 10^3]
Beamforming	0.96	1.09	1.10	1.25
Capon Beamforming	0.98	1.35	1.38	1.58
MUSIC	0.66	0.66	0.68	0.69
CS	0.71	0.71	0.72	0.72
3-D inversion	**0.57**	**0.58**	**0.60**	**0.60**

TABLE II

EVOLUTION OF THE MACT VALUE WHEN THE SCATTERER AMPLITUDE IS FIXED UNIFORMLY FOR EACH AZIMUTH. THE EXPERIENCE IS REPEATED FOR INCREASING AMPLITUDE INTERVAL SIZE.
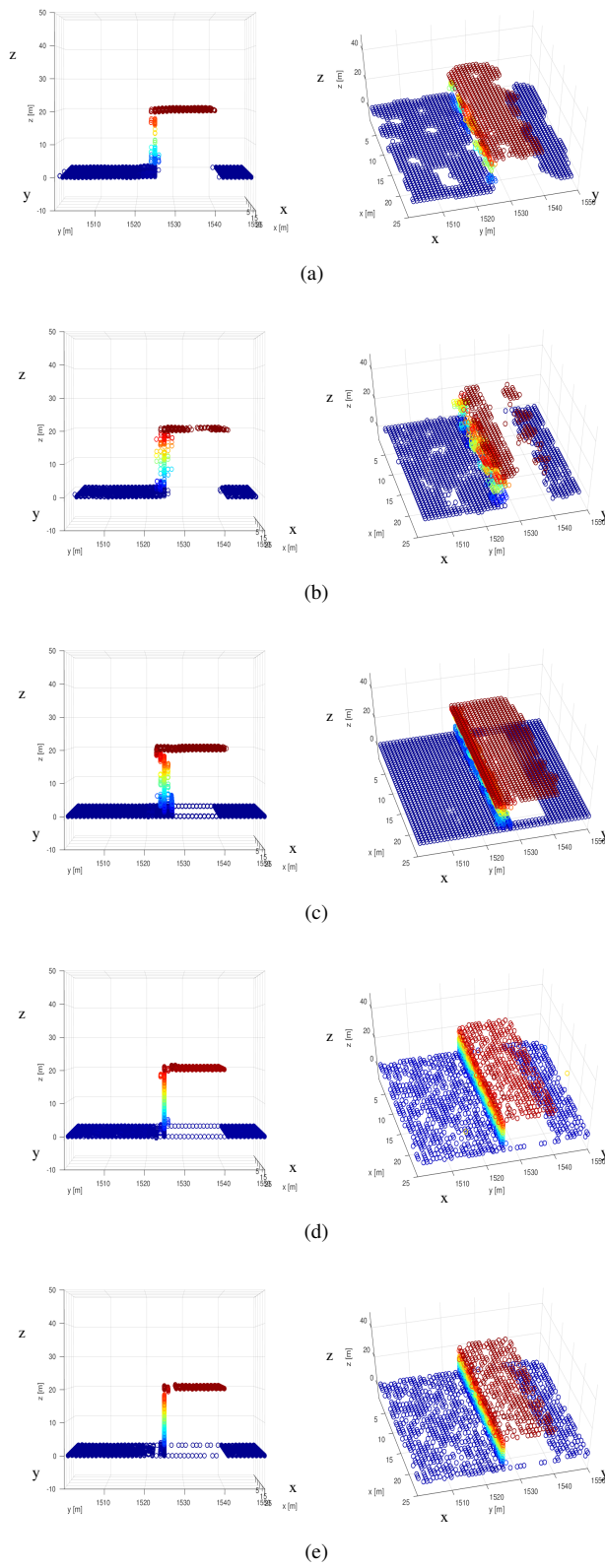
Fig. 13. Best representation on the `Reg` scene using classical beamforming (a), Capon beamforming (b), MUSIC (c), CS (d) and our tomographic 3D inversion (e).
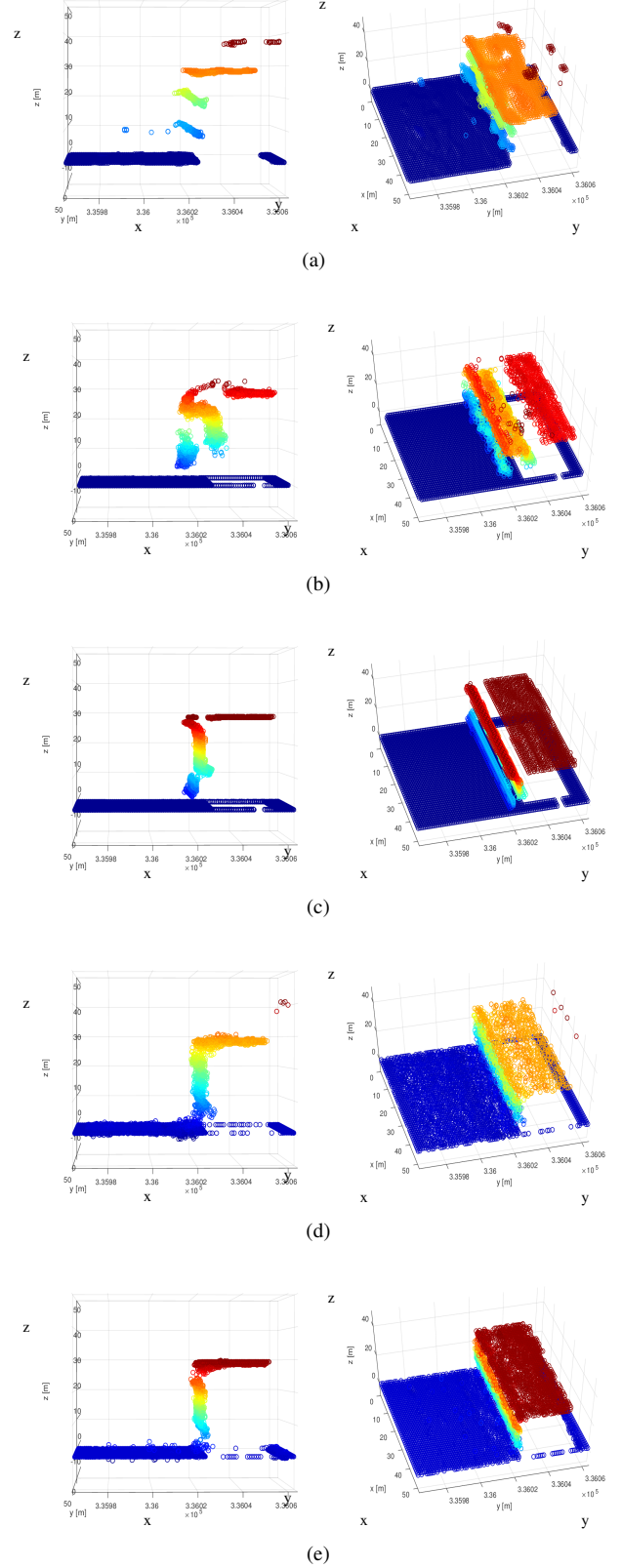

Fig. 14. Best representation on the `TSX` scene using classical beamforming (a), Capon beamforming (b), MUSIC (c), CS (d) and our tomographic 3D inversion (e).

E. Real Data

We test our algorithm and compare it to the other SAR tomographic estimators on a stack of 40 TerraSAR-X spotlight images acquired over the *front de Seine* in the south-west of Paris, France. The slant-range resolution is 0.45 m and the

(a)

(b)

Fig. 15. Accuracy vs. completeness for classical beamforming, Capon beamforming, MUSIC, Compress Sensing and our 3D Inversion algorithm. In (a) the evaluation is done on the `Reg` scene 6 and in (b) on the `TSX` scene 6.

azimuth resolution 0.87 m. The observed scene is presented in Fig. 16[3]. The spatial and temporal baselines are shown in Fig. 17. The total spatial baseline Δb span is more than 775 m and the total temporal baseline more than 5 years with a large gap of almost two years. The theoretical resolutions in h and z are given by :

$$\delta_z = \delta_h \sin(\theta) = \frac{\lambda r \sin(\theta)}{2\Delta b} = 6.99m \qquad (21)$$

with an incidence angle $\theta = 0.6$ rad and wavelength $\lambda = 0.0311$ m. Sparse reconstructions in urban SAR tomography have been shown to significantly improve this resolution [37].

To evaluate the different tomographic estimators we use a rough ground truth of the scene. The ground truth surface is shown in Fig. 18[4]. The big structures such as the skyscrapers or the tall buildings are well represented by smooth polygons. However, the vegetation and the small structures on the docks, bridges or streets are not represented. We therefore performed the evaluation only on some areas where we found the ground truth to be accurate enough. The results of the evaluation are given in Fig. 18.

It is noteworthy that MUSIC performs particularly well on areas containing a lot of ground surface. In the first test, the sub-scene is composed of a mixture of tall buildings,
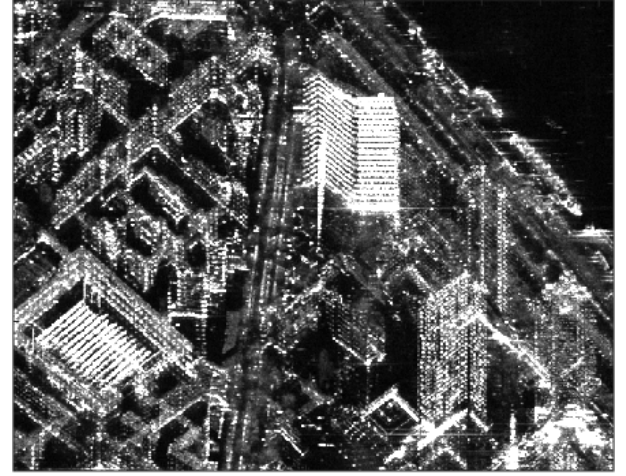

(a)

(b)

Fig. 16. (a) Image of the temporal mean of all the iamges in the stack and (b) the corresponding optic image obtained from Geoportail (IGN).The building back-scattering a strong signal in the center af the image is the Mirabeau Tower. In the bottom left of the image, the structure presenting a strong periodic back-scattering signal is the Ministry of Foreign Affair.
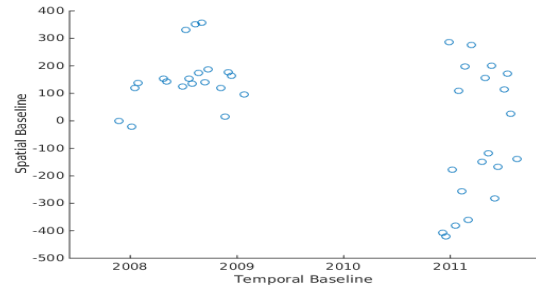

Fig. 17. Orthogonal spatial baseline vs. temporal baseline for all the 40 acquisitions.

---

[3]These data have been provided by the DLR in the framework of the project LAN1746.

[4]Ground truth on Paris complimentary provided by the IGN, France.

small structures near the ground level and smooth flat areas

(streets and docks). As the diffuse signal back-scattered by flat surfaces is much weaker than the one reflected by the dihedral or trihedral structures, it is generally not taken into account by sparse representations of the scene. CS is then unable to correctly represent ground areas. Taking more points into account results only in more outliers generally due to the sidelobes of the brightest points. The averaging step used to estimate the covariance matrix allows MUSIC to retrieve part of the ground points or to extend the signal coming from one punctual target close to the ground to neighboring pixels. Unlike the non-parametric beamforming methods, MUSIC is designed to retrieve a sparse scene which removes most of the side lobes and outliers. MUSIC is then able to outperform CS according to our evaluation method that includes scatterers on the ground. On the second tested area corresponding to the red rectangle in Fig. 18, the performances of MUSIC and CS in terms of accuracy are very similar. MUSIC seems to perform a little bit better than CS. However, this may again be due to its ability to retrieve more points on the ground and the small gap between the two methods needs to be taken with caution as the used ground truth is coarse.

In both experiments, the presented algorithm achieves the best scores in term of accuracy and completeness. Moreover, the analysis of the metrics on real data shows that in all the possible ways of selecting the reconstructed voxels, our tomographic 3-D inversion algorithm performs a better reconstruction than the compared estimators. For this scene the parameters are $\mu_{\ell_1} = 12$, $\mu_z = 300$, $\mu_y = 100$ and $\mu_x = 100$ and were computed from the TSX simulations.

The 3-D plain representation of the scene obtained by our approach is presented in Fig. 21. The heat color is chosen to be proportional to the intensity of the voxels normalized by the averaged intensity of the corresponding radar cell. This normalization is used only to help the visualization as the dynamic range of SAR images is very high. In Fig. 22, the same image is shown except that the near zero intensity voxels are made fully transparent and the ground truth is superimposed. The results obtained on the same scene with conventional CS are shown in Fig. 19 and Fig. 20. The colormap are the same for both representations and are chosen to have the dark blue color associated with intensities below 10^{-4} times the maximum voxel amplitude. This visually set to zero dark voxels that still have a non-zero intensity as the algorithm may not have fully converged. An other way to have a sparser representation of the modulus of the signal is to display the image w at convergence of the algorithm. In the 3-D representation, most of the isolated outliers are suppressed thanks to the spatial smoothing and the buildings where most of the high reflectivity voxels are massed are retrieved. The smoothing denoising however comes with the cost of slightly blurred scatterers. We can see that most of the very bright buildings present in the SAR images in Fig. 16 are well reconstructed. The two towers that were mostly missing in the 2D intensity image have a point-like representation but are fairly visible in this 3-D rendering.

By using the georeferencing of the voxels, we can project the corresponding point cloud into Google Earth ©. The results are presented in Fig. 23. This step allows us to have more details in the visualization of the scene. We can now see that some points above the Mirabeau Tower are relevant: they correspond to the structure of its rooftop. Our reconstruction method also correctly identified parts of the structures on the dock and a large amount of buildings hidden behind the Mirabeau tower.

V. CONCLUSION

In this paper a new tomographic 3-D inversion method has been proposed. This method is shown to be effective to inverse a SAR tomographic stack in a dense urban area. The chosen spatial regularization strategy takes advantage of both the natural sparsity of targets in urban environments and the spatial smoothness (i.e., spatial proximity of scatterers). Our approach generalizes compressed sensing techniques and includes, in an adaptive way, a spatial smoothing.

To be able to compare our method with other estimators used in SAR tomography we extended the protocol of evaluation based on accuracy and completeness proposed in [19]. The tests done on numerical simulations and on a stack of TerraSAR-X images in a urban area showed that our method leads to better results than existing approaches in terms of accuracy / completeness compromise.

This 3D inversion algorithm opens a door to more sophisticated approaches taking into account different priors describing the data. Indeed, the main structure of the tomographic 3D inversion would remain unchanged if other regularization terms were considered within a variable splitting approach. The research of appropriate priors to scenes with a volumetric density of reflectivity such as ice or forest will be the subject of further works.

VI. ACKNOWLEDGEMENTS
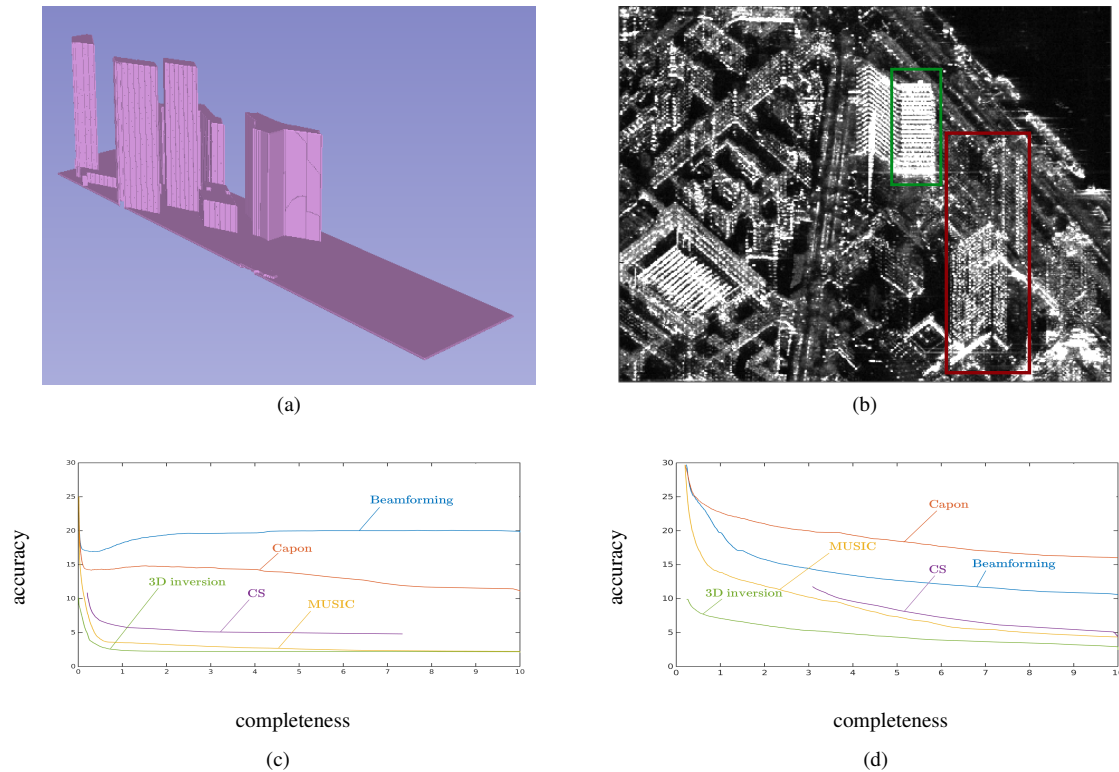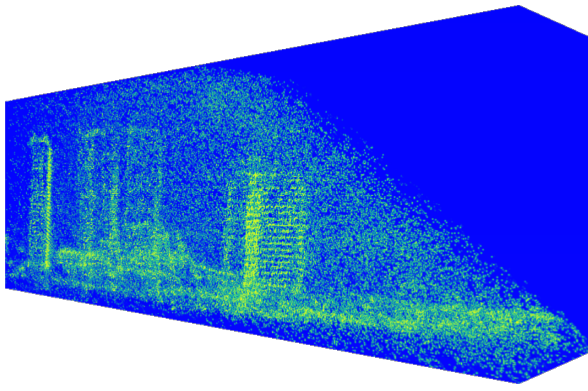
(a)

(b)

completeness

(c)

completeness

(d)

Fig. 18. Scores accuracy versus completeness based on the ground truth in (a) for the tomographic estimators classical beamforming, Capon beamforming, MUSIC, CS and the proposed 3D inversion. In (c) the evaluation is done over the area corresponding to the red rectangle in (b). In (d) the test area is inside the green rectangle.
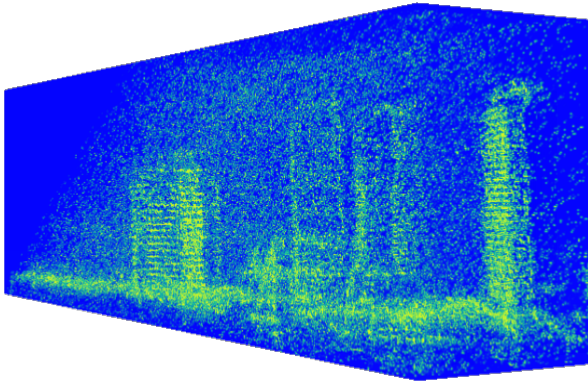
REFERENCES

[1] Y. Huang, L. Ferro-Famil, and A. Reigber, "Under foliage object imaging using SAR tomography and polarimetric spectral estimators," in *8th European Conference on Synthetic Aperture Radar*, June 2010, pp. 1–4.

[2] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2142–2152, Sep 2000.

[3] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, March 2013.

[4] R. M. P. Stoica, in *Introduction to Spectral Analysis*. Prentice Hall, 1997.

[5] F. Gini and F. Lombardini, "Multibaseline cross-track SAR interferometry: a signal processing perspective," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 71–93, Aug 2005.

[6] S. Tebaldini and F. Rocca, "Multibaseline Polarimetric SAR Tomography of a Boreal Forest at P- and L-Bands," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 232–246, Jan 2012.

[7] C. A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger, "NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2021–2038, April 2015.

[8] C. A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting Patch Similarity for SAR Image Processing: The nonlocal paradigm," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 69–78, July 2014.

[9] C. Rambour, L. Denis, F. Tupin, J.-M. Nicolas, H. Oriot, L. Ferro-Famil, and C. Deledalle, "Similarity Criterion for SAR Tomography over Dense Urban Area," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2017.

[10] S. Tebaldini, T. Nagler, H. Rott, and A. Heilig, "Imaging the internal structure of an alpine glacier via l-band airborne sar tomography," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7197–7209, Dec 2016.

[11] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-Dimensional SAR Focusing From Multipass Signals Using Compressive Sampling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 488–499, Jan 2011.

[12] X. X. Zhu and R. Bamler, "Tomographic SAR Inversion by L_1 - Norm Regularization ; The Compressive Sensing Approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, Oct 2010.

[13] ——, "Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 247–258, Jan 2012.

[14] G. Fornaro, A. Pauciullo, D. Reale, M. Weiß, A. Budillon, and G. Schirinzi, "Compressive sensing and generalized likelihood ratio test in SAR tomography," in *2016 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, Sept 2016, pp. 90–94.

[15] A. Budillon, A. C. Johnsy, and G. Schirinzi, "A fast support detector for superresolution localization of multiple scatterers in sar tomography," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 6, pp. 2768–2779, June 2017.

[16] X. X. Zhu and M. Shahzad, "Facade Reconstruction Using Multiview Spaceborne TomoSAR Point Clouds," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3541–3552, June 2014.

[17] A. Ley, O. D'Hondt, and O. Hellwich, "Regularization and Completion of TomoSAR Point Clouds in a Projected Height Map Domain," pp. 1–11, 04 2018.

[18] O. D'Hondt. and Al., "Impact of non-local filtering on 3D reconstruction from tomographic SAR data," in *2017 IEEE International Geoscience and Remote Sensing Symposium*, 2017.

[19] O. D'Hondt, C. López-Martínez, S. Guillaso, and O. Hellwich, "Nonlocal Filtering Applied to 3-D Reconstruction of Tomographic SAR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 272–285, Jan 2018.

[20] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 3, pp. 507–517, March 2003.

[21] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 8–20, Jan 2001.

[22] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57(8), pp. 1408–1418, 1969.

[23] M. Porfiri, L. Ferro-Famil, and J.-M. Nicolas, "Building profile reconstruction using TerraSAR-X data time-series and tomographic techniques," in *Analysis of Multitemporal Remote Sensing Images (Multi-Temp), 2015 8th International Workshop on the*, Annecy, France, 2015, pp. 1–4. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01279745

[24] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar 1986.

[25] S. Tebaldini, F. Rocca, and A. M. Guarnieri, "Model Based SAR Tomography of Forested Areas," in *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, July 2008, pp. II–593–II–596.

[26] M. Schmitt and X. X. Zhu, "Demonstration of Single-Pass Millimeterwave SAR Tomography for Forest Volumes," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 202–206, Feb 2016.

[27] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2436–2449, Nov 1991.

[28] D. Reale, G. Fornaro, A. Pauciullo, X. Zhu, and R. Bamler, "Tomographic Imaging and Monitoring of Buildings With Very High Resolution SAR Data," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 661–665, July 2011.

[29] X. X. Zhu and R. Bamler, "Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 247–258, Jan 2012.

[30] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[31] R. Mourya, L. Denis, J.-M. Becker, and É. Thiébaut, "Augmented Lagrangian without alternating directions: Practical algorithms for inverse problems in imaging," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1205–1209.

[32] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[33] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Transactions on Mathematical Software (TOMS)*, vol. 23, no. 4, pp. 550–560, 1997.

[34] E. Thiébaut, "Optimization issues in blind deconvolution algorithms," in *Astronomical Data Analysis II*, vol. 4847. International Society for Optics and Photonics, 2002, pp. 174–184.

[35] A. Budillon, A. C. Johnsy, and G. Schirinzi, "Extension of a Fast GLRT Algorithm to 5D SAR Tomography of Urban Areas," *Remote Sensing*, vol. 9, no. 8, p. 844, 2017. [Online]. Available: https://doi.org/10.3390/rs9080844

[36] B. S. He, H. Yang, and S. L. Wang, "Alternating Direction Method with Self-Adaptive Penalty Parameters for Monotone Variational Inequalities," *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 337–356, Aug 2000. [Online]. Available: https://doi.org/10.1023/A:1004603514434

[37] X. X. Zhu and R. Bamler, "Very High Resolution Spaceborne SAR Tomography in Urban Environment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4296–4308, Dec 2010.

[38] Andriy Fedorov et al., "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323 – 1341, 2012, quantitative Imaging in Cancer. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0730725X12001816
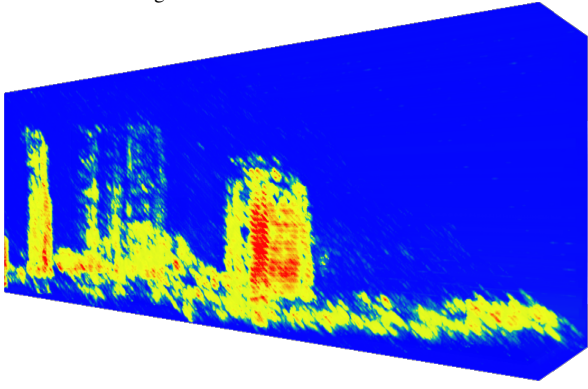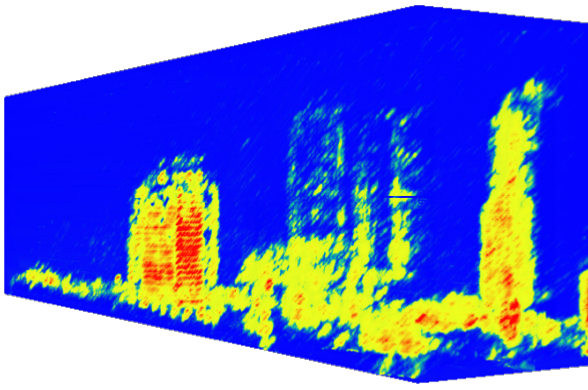
(a)

(b)

Fig. 19.   Visualization of the module of the reconstructed cube [38] from a stack of 40 TerraSAR-X images of Paris (cf. Fig.16) using the conventional CS algorithm.
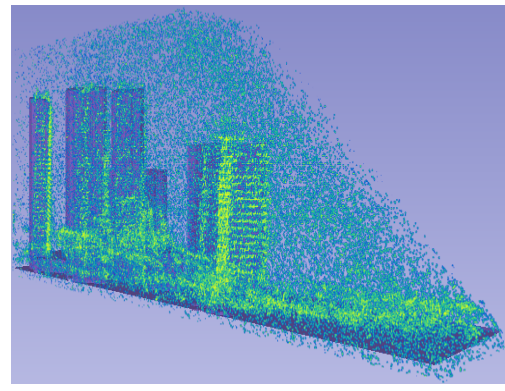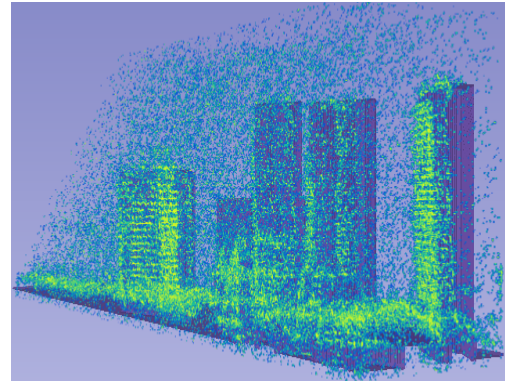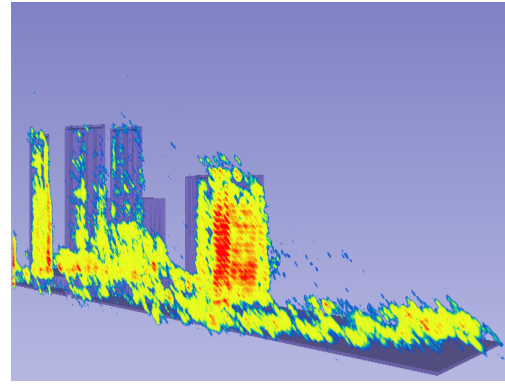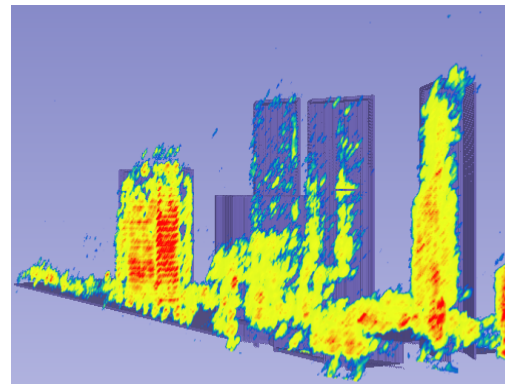

(a)

(b)

Fig. 20.   Visualization of the thresholded module of the reconstructed cube [38] using the conventional CS algorithm with the ground truth.

(a)

(b)

Fig. 21.   Visualization of the module of the reconstructed cube [38] from a stack of 40 TerraSAR-X images of Paris (cf. Fig.16) using the 3D inversion algorithm.
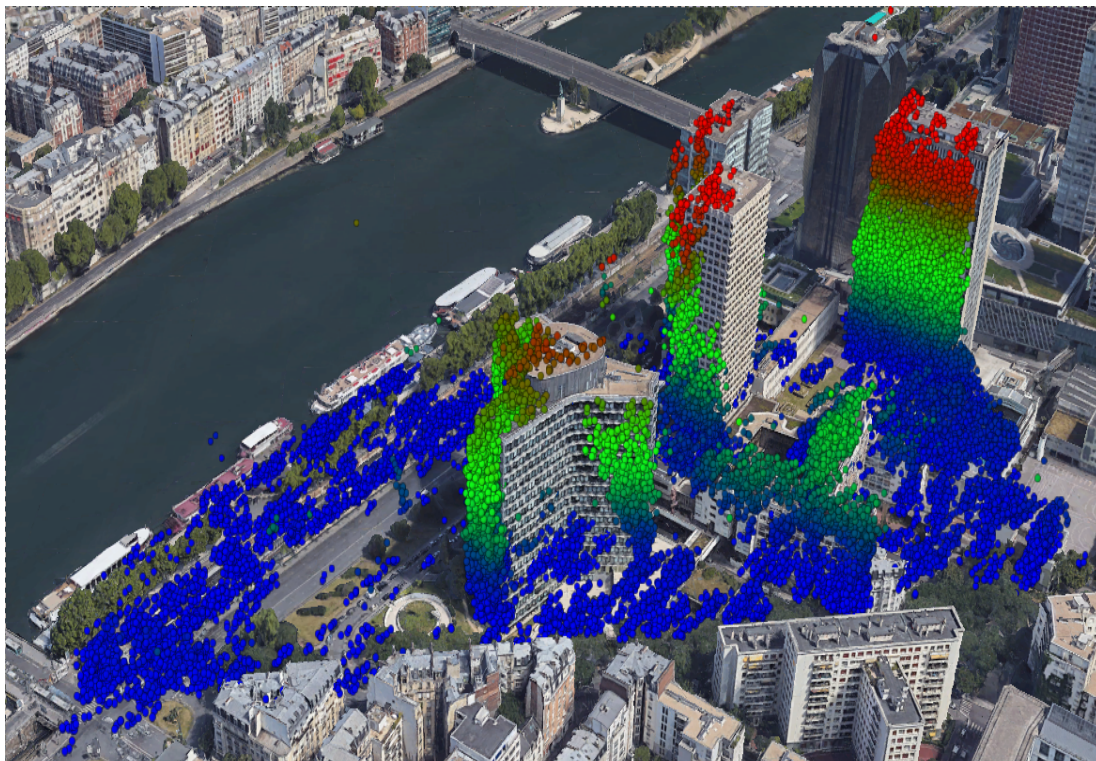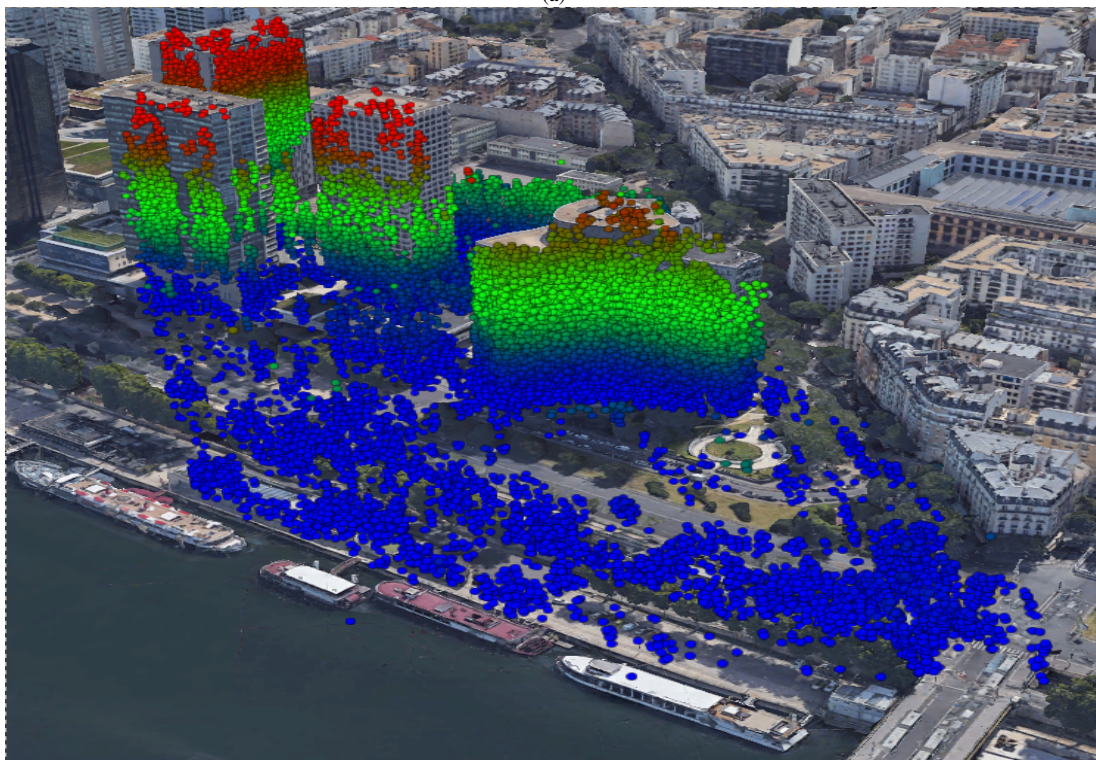

(a)

(b)

Fig. 22.   Visualization of the thresholded module of the reconstructed cube [38] using the 3D inversion algorithm with the ground truth.

(a)

(b)

Fig. 23.   Visualization of the point cloud obtained from the reconstructed cube with projected in Google Hearth © .