

# Statistical Regularization for Enhanced TomoSAR Imaging

Gustavo Martín del Campo, Matteo Nannini and Andreas Reigber

**Abstract**—One of the main topics in synthetic aperture radar (SAR) tomography (TomoSAR) is the estimation of the vertical structures' location, which scatter the field back towards the sensor, constrained to a reduced number of passes. Moreover, the introduction of artifacts and the increase of the ambiguity levels due to irregular sampling, consequence of non-uniform acquisition constellations, complicate the accurate estimation of the source parameters. Pursuing the alleviation of such drawbacks, the use of statistical regularization approaches, based on the maximum-likelihood (ML) estimation theory, has been successfully demonstrated in the previous related studies. However, these techniques are constrained to the assumption that the probability density function (pdf) of the observed data is Gaussian. In this paper, in order to solve the ill-posed non-linear TomoSAR inverse problem, we relax this assumption and apply the weighted covariance fitting (WCF) criterion instead. The latter alleviates the previously mentioned drawbacks, and retrieves a power spectrum pattern (PSP) with an outline more similar to the expected one, i.e., recovered using matched spatial filtering (MSF) with a higher number of tracks. First, we present the mathematical background of the related regularization methods, adapted to solve the TomoSAR inverse problem, from which we derive our novel technique, named WISE (WCF-based Iterative Spectral Estimator). Then, the differences and similarities between the addressed regularization approaches are discussed, besides their main advantages and disadvantages. Finally, the implementation details of WISE are treated, along with simulated examples and experimental results gotten from a forested test site.

**Index Terms**—Maximum-likelihood (ML), power spectrum pattern (PSP), spectral analysis (SA), synthetic aperture radar (SAR) tomography (TomoSAR), weighted covariance fitting (WCF).

## LIST OF ACRONYMS

BMR	Bayes minimum risk
CLS	Constrained least squares
CS	Compressed sensing
DOA	Direction of arrival
ESPRIT	Estimation of signal parameters via rotational invariance techniques
IAA-ML	Iterative adaptive approach for ML
LS	Least squares
MAP	Maximum a posteriori probability
MARIA	ML-inspired adaptive robust iterative approach
ML	Maximum-likelihood
MSE	Mean square error

MSF	Matched spatial filter
MUSIC	Multiple signal classification
PLOS	Perpendicular to the line-of-sight
POCS	Projector onto convex solution sets
PSP	Power spectrum pattern
SA	Spectral analysis
SAR	Synthetic aperture radar
SM	Scattering mechanism
SPICE	Semiparametric/sparse iterative covariance-based estimation
TomoSAR	SAR tomography
WAVAB	WDT-refined virtual adaptive beamforming
WDT	Wavelet domain thresholding
WCLS	Weighted CLS
WCF	Weighted covariance fitting
WISE	WCF-based iterative spectral estimator

## GLOSSARY OF NOTATION

$\langle \cdot \rangle$	Averaging operator
*	Conjugate
$\mathbf{D}(\mathbf{u})$	Diagonal matrix with vector $\mathbf{u}$ at the principal diagonal
$\ \cdot\ $	Euclidean $\ell_2$ -norm
$\mathbf{E}(\cdot)$	Expectation operator
+	Hermitian conjugate (adjoint)
$\mathbf{I}$	Identity matrix
$[\mathbf{u}, \mathbf{v}]$	Inner product between vector $\mathbf{u}$ and vector $\mathbf{v}$
$\{\mathbf{V}\}_{\text{diag}}$	Main diagonal of matrix $\mathbf{V}$
$\mathbf{T}$	Transpose
$\text{tr}\{\mathbf{V}\}$	Trace of matrix $\mathbf{V}$

## I. INTRODUCTION

**S**YNTHETIC aperture radar (SAR) tomography (TomoSAR) has the main goal of estimating the location of the vertical structures that scatter the field back towards the sensor [1] – [3]. A stack of coregistered imagery is acquired from several tracks with different lines-of-sight (as shown in Fig. 1), in order to be coherently combined using SAR interferometric techniques. Afterward, the vertical profiles that characterize the observed backscattered power spectrum pattern (PSP), for each azimuth-range resolution cell, are recovered using spectral analysis (SA) based methods [1] – [7].

For practical TomoSAR sensing scenarios, the number of utilized tracks is constrained to a reduced number [4], [5] i.e., in order to avoid temporal decorrelation issues due to the

The authors are with the Microwaves and Radar Institute, German Aerospace Center (DLR), 82230 Oberpfaffenhofen, Germany (e-mail: [gustavo.martindelcampobecerra@dlr.de](mailto:gustavo.martindelcampobecerra@dlr.de); [matteo.nannini@dlr.de](mailto:matteo.nannini@dlr.de); [andreas.reigber@dlr.de](mailto:andreas.reigber@dlr.de))

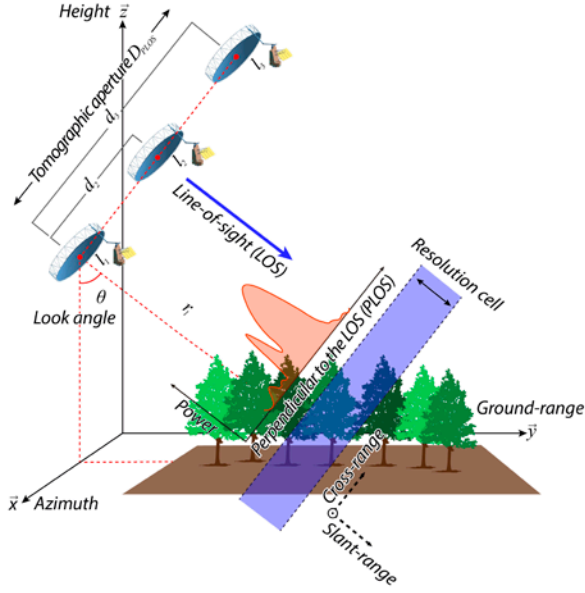


Fig. 1. TomoSAR acquisition geometry using parallel passes (not to scale).

revisit time. Moreover, non-uniform acquisition constellations and irregular sampling introduce artifacts and increase ambiguity [4] – [7]. All these drawbacks complicate the accurate estimation of the source parameters.

Conventional approaches to solve the ill-conditioned TomoSAR non-linear inverse problem retrieve an estimate of the PSP, in the direction perpendicular to the line-of-sight (PLOS), applying SA-inspired techniques in the context of the direction of arrival (DOA) estimation theory [3], [7], [8].

Among the DOA non-parametric methods, there are the matched spatial filter (MSF) [3] and Capon beamforming [7] – [9] techniques. MSF is well suited to cope with scenarios characterized by the presence of distributed scatterers, since it preserves radiometric accuracy; whereas Capon beamforming trades radiometric accuracy for enhanced resolution. Yet, ultimately, the resolution capability of these SA-based estimators highly depends on the total baseline span [3], [4], [7], the so-called tomographic aperture (see Fig. 1). Meaning, that the arbitrarily closely spaced phase-centers may not be distinguished owed to the limited resolution. For the aim of this work, we denote as phase-centers to the mean heights of the multiple backscattering sources.

The parametric DOA estimation methods include super-resolution approaches as multiple signal classification (MUSIC) [3], [4], [6] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [8]. Both, MUSIC and ESPRIT, improve considerably the resolution in the PLOS direction and mitigate the effect of sidelobes. Nevertheless, these methods have the main drawback related to the assumption that the scene is composed by a known finite number of point-type like backscattering sources.

Although point-type responses are not common in real-world TomoSAR operating scenarios, the introduction of distributed models to retrieve the approximate number of sources that define the signal-subspace [4], [6], [7], has demonstrated the successful implementation of the

aforementioned parametric DOA-based estimation techniques for TomoSAR applications.

On the other hand, taking advantage of the sparse representations of the cross-range profiles in the wavelet domain, super-resolved TomoSAR-adapted compressed sensing (CS) based approaches are introduced in [10], [11]. However, these CS-based techniques imply a considerable computational burden, due to their iterative nature and due to the non-availability of adapted efficient convex optimization algorithms [11, Section V].

As an alternative to the aforementioned commonly performed TomoSAR-adapted focusing methods, this paper tackles the ill-posed non-linear TomoSAR inverse problem using statistical regularization. In the previous related studies [12], [13], the use of different regularization approaches, based on maximum-likelihood (ML), has been successfully verified. Nonetheless, these techniques are constrained to the assumption that the probability density function (pdf) of the observed data is Gaussian. In this article, we relax this assumption and introduce a novel technique based on the weighted covariance fitting (WCF) criterion, called WISE (WCF-based iterative spectral estimator). This method ensures the alleviation of the previously mentioned drawbacks and retrieves a PSP with an outline more similar to the expected one, i.e., recovered using MSF with a higher number of tracks.

First, the TomoSAR signal model at hand is presented; then, we address the different related regularization approaches from which WISE evolved. These techniques are assembled into two main groups: (i) deterministic descriptive regularization techniques, inspired by the Tikhonov's regularization theory; and (ii) statistical regularization methods, in the context of the statistical decision-making theory. Later on, the differences and similarities between each regularization approach are discussed, as well as their main advantages and disadvantages. Finally, the implementation details of the proposed technique are treated, along with simulated examples and experimental results gotten from a forested test site.

This work is organized in the following manner. The TomoSAR signal model is stated in Section II. Section III addresses the related TomoSAR-adapted deterministic descriptive regularization approaches. Section IV presents the related TomoSAR-adapted statistical regularization techniques and introduces the novel WCF solver. Section V offers a comparison between each addressed regularization approach. Section VI presents the implementation details of the introduced WISE regularization technique. Simulated examples and experimental results are treated in Section VII and Section VIII respectively. Finally, the concluding remarks are addressed in Section IX.

## II. TOMOSAR SIGNAL MODEL

The Fourier spectrum of the observed data over the PLOS cross-range axis, in a particular azimuth-range resolution cell, can be modelled as an ensemble of some dominating scattering mechanisms (SMs) [14] – [16]. In this paper, we assume no a priori knowledge about the number, location or

distribution of these SMs. Thus, we resort to a non-parametric SMs model-free approach, in which the scatterers are assumed to be distributed over unknown support regions along the sensing search area.

Within the context of the DOA estimation theory [3], [7], [8, Chapter 6], consider a TomoSAR acquisition constellation, composed by  $L$  tracks, as a linear array. Assuming coregistration independent on height, the discrete-form representation of the data signal vector  $\mathbf{y}$  represents the set of  $L$  focused signals for a given azimuth-range position, and is related to the complex random scene reflectivity vector  $\mathbf{s}$  via the linear equation of observation [13], [17] – [19],

$$\begin{bmatrix} \mathbf{y} \end{bmatrix}_{L \times I} = \begin{bmatrix} \mathbf{A} \end{bmatrix}_{L \times M} \begin{bmatrix} \mathbf{s} \end{bmatrix}_{M \times I} + \begin{bmatrix} \mathbf{n} \end{bmatrix}_{L \times I}. \quad (1)$$

Vector  $\mathbf{s}$  gathers a sufficiently large amount of samples  $M$ , which characterize the continuous complex random reflectivity for a certain azimuth-range location. The  $M$  samples are taken from the elevation positions  $\{z_m\}_{m=1}^M$  displaced along the PLOS height direction. The  $L \times M$  steering matrix  $\mathbf{A}$  is the signal formation operator that maps  $S \rightarrow Y$ , the source Hilbert signal space  $S$  onto the observation Hilbert signal space  $Y$ . Finally, vector  $\mathbf{n}$  accounts for the additive noise.

The steering matrix  $\mathbf{A}$  is composed of  $M$   $L$ -dimensional steering vectors  $\{\mathbf{a}_m\}_{m=1}^M$ , which contains the interferometric phase information associated to a source located at the PLOS elevation position  $\{z_m\}_{m=1}^M$ , above the reference focusing plane, which, for a particular elevation position  $z$  is given by [7],

$$\mathbf{a}(z) = \left[ 1, \exp\{jk_{z_2}z\}, \dots, \exp\{jk_{z_L}z\} \right]^T, \quad (2)$$

in which  $\{k_{z_l}\}_{l=2}^L$  is the two-way vertical wavenumber between the master track and the  $l$ -th acquisition positions,

$$k_{z_l} = \left( \frac{4\pi}{\lambda} \right) \left( \frac{d_l}{r_1 \sin \beta} \right), \quad l = 2, \dots, L; \quad (3)$$

where  $\lambda$  is the carrier wavelength,  $\beta$  stands for the incidence angle,  $r_1$  is the slant-range distance between the master track and the target, and  $\{d_l\}_{l=2}^L$  is the cross-range oriented baseline distance between the master position and the  $l$ -th acquisition position.

Vectors  $\mathbf{s}$ ,  $\mathbf{n}$ ,  $\mathbf{y}$ , represent complex random Gaussian zero-mean vectors composed of the decomposition coefficients  $\{s_m\}_{m=1}^M$ ,  $\{n_l\}_{l=1}^L$  and  $\{y_l\}_{l=1}^L$ , of the corresponding continuous

signal  $S$ , noise  $N$  and observation  $Y$  fields, respectively [17] – [19]. These vectors are specified by the employed TomoSAR modulation and acquisition formats and are characterized by the correlation matrices [7]

$$\mathbf{R}_s = \mathbf{D}(\mathbf{b}) = \text{diag}(\mathbf{b}), \quad (4)$$

$$\mathbf{R}_n = N_0 \mathbf{I}, \quad (5)$$

and

$$\mathbf{R}_y = \mathbf{A} \mathbf{R}_s \mathbf{A}^+ + \mathbf{R}_n, \quad (6)$$

correspondingly, where  $^+$  stands for the adjoint operator (Hermitian conjugate) and  $N_0$  is the power spectral density of the white noise power [3]. Vector  $\mathbf{b}$  at the principal diagonal of the diagonal matrix  $\mathbf{D}(\mathbf{b})$ , composed of the averaged entries

$$\{b_m = \langle |s_m|^2 \rangle\}_{m=1}^M, \quad (7)$$

at each  $m$ -th PLOS elevation position  $\{z_m\}_{m=1}^M$ , defines the backscattering power, referred to also as the PSP, i.e., the second-order statistics of the complex reflectivity vector  $\mathbf{s}$ .

We refer to the unconditional model in [7], well adapted for the representation of natural environments, in which the distributed media is characterized by a scattering response with a random behavior, i.e., rough surfaces and volumetric targets as forests. Therefore, in equations (4), (6), we only consider the unconditional model contributions as described in [7, Eq. 10]. The correlation matrix  $\mathbf{R}_s$  is modelled then as a diagonal matrix, since it is assumed that the entries of vector  $\mathbf{s}$  are uncorrelated.

The TomoSAR inverse problem is stated as follows: given the complex data recordings  $\mathbf{y}$ , defined by the linear equation of observation in (1), using the signal formation operator (steering matrix)  $\mathbf{A}$ , and some a priori information about the statistics of the signal and noise, derive a solution operator  $\mathbf{F}$ , which when applied to the observed TomoSAR data  $\mathbf{y}$ , produces an estimate of the original signal  $\hat{\mathbf{s}} = \mathbf{F} \mathbf{y}$ , being optimal/sub-optimal in some statistical or non-statistical sense.

Yet, this article treats the non-linear case of the stated TomoSAR inverse problem, which consists in retrieving an estimate  $\hat{\mathbf{b}}$  of the PSP vector  $\mathbf{b}$  in (7), expressed in the following manner,

$$\hat{\mathbf{b}} = \left\{ \mathbf{F} \mathbf{y} \mathbf{y}^+ \mathbf{F}^+ \right\}_{\text{diag}}, \quad (8)$$

where operator  $\{\cdot\}_{\text{diag}}$  retrieves a vector composed of the elements at the main diagonal of the embraced matrix.

Thus, the TomoSAR problem at hand entails the reconstruction of an optimal/sub-optimal (in some statistical or non-statistical sense) estimate of the vertical profiles that characterize the illuminated scene PSP, within each observed azimuth-range resolution cell. Notice that, to handle multiple non-deterministic sources and in order to increase accuracy in presence of signal-dependent (multiplicative) noise [3], the addressed solver in (8) may better employ the measurement second-order statistics presented by the data covariance matrix [3], [4], [7],

$$\mathbf{Y} = \frac{1}{J} \sum_{j=1}^J \mathbf{y}_{(j)} \mathbf{y}_{(j)}^+, \quad (9)$$

where  $j = 1, \dots, J$  indicates one of  $J$  independent realizations (looks, snapshots) of the TomoSAR signal acquisitions. Technically, multi-looking is accomplished through the averaging of adjacent values or from multiple reduced resolution focalizations of the cell of interest via azimuth sub-apertures.

By instance, if we take into account the classical MSF beamforming technique, given in a matrix-vector form by [19]

$$\hat{\mathbf{b}}_{\text{MSF}} = \left\{ \mathbf{A}^+ \mathbf{Y} \mathbf{A} \right\}_{\text{diag}}, \quad (10)$$

then, we can realize that the solution operator is defined by  $\mathbf{F} = \mathbf{A}^+$ .

Normally  $L < M$ , this makes the TomoSAR problem ill-posed in the Hadamard sense [20], [21], since the *uniqueness* condition is not accomplished. There are more unknowns  $M$  than equations  $L$  in (1), meaning that an infinite number of solutions to the equation of observation exists. Moreover, the unavoidable presence of additive and signal-dependent noise in the observed data  $\mathbf{y}$ , along with the inherent imprecisions of any problem model, adds statistical uncertainty to the PSP estimation problem.

At this stage of our analysis, we may deduce that, some form of constraints should be imposed to the solution operator  $\mathbf{F}$ , in order to guarantee well-conditioned solutions to the TomoSAR inverse problem. A problem is considered well-posed in the Hadamard sense [20], [21], when three conditions are satisfied: (i) a solution exists; (ii) the solution is unique; (iii) the solution varies continuously with the data.

### III. DETERMINISTIC DESCRIPTIVE REGULARIZATION

The descriptive idea of the Tikhonov's (deterministic) regularization is to replace the ill-posed TomoSAR problem with a well-conditioned optimization problem.

#### A. Constrained Least Squares (CLS)

For the case of the constrained least squares (CLS) approach, the objective function  $\Lambda(\mathbf{s})$  refers to the standard least squares (LS) fitting criterion augmented with the minimum norm cost function [22] – [24], yielding

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \left\{ \|\mathbf{y} - \mathbf{A}\mathbf{s}\|^2 + \alpha \|\mathbf{s}\|^2 \right\}, \quad (11)$$

where  $\alpha$  is a regularization parameter that acts as an adjustable degree-of-freedom.

The expansion of the CLS objective function  $\Lambda(\mathbf{s})$  in terms of inner products, i.e.,  $\|\mathbf{x}\|^2 = [\mathbf{x}, \mathbf{x}] = \mathbf{x}^+ \mathbf{x}$ , is expressed as

$$\Lambda(\mathbf{s}) = \mathbf{s}^+ \mathbf{A}^+ \mathbf{A} \mathbf{s} - \mathbf{y}^+ \mathbf{A} \mathbf{s} - \mathbf{s}^+ \mathbf{A}^+ \mathbf{y} + \alpha (\mathbf{s}^+ \mathbf{s}), \quad (12)$$

where the terms independent to vector  $\mathbf{s}$  have been ignored. Next, we apply the standard gradient method to solve (11) with respect to  $\mathbf{s}$  (the employed differentiation rules are gathered in the Appendix), resulting in the derivative expression,

$$\frac{\partial}{\partial \mathbf{s}} \Lambda(\mathbf{s}) = -\mathbf{A}^+ (\mathbf{y} - \mathbf{A} \mathbf{s}) + \alpha (\mathbf{s}) = \mathbf{0}. \quad (13)$$

Rearranging the terms in (13),

$$(\mathbf{A}^+ \mathbf{A} + \alpha \mathbf{I}) \mathbf{s} = \mathbf{A}^+ \mathbf{y}, \quad (14)$$

infers the desired CLS estimator  $\hat{\mathbf{s}} = \mathbf{F}_{\text{CLS}} \mathbf{y}$ , with the solution operator

$$\mathbf{F}_{\text{CLS}} = (\mathbf{A}^+ \mathbf{A} + \alpha \mathbf{I})^{-1} \mathbf{A}^+; \quad (15)$$

which admits the usage of  $\mathbf{F}_{\text{CLS}}$  in (8) to retrieve an estimate  $\hat{\mathbf{b}}_{\text{CLS}}$  of the PSP.

The regularization term  $\alpha \mathbf{I}$  in (15) is necessary to make the solution well-conditioned. Thus, solving (8) requires the choice of a suitable value  $\alpha > 0$ , which in the limiting case when  $\alpha \rightarrow 0$ , the problem is reduced to its unconstrained ill-posed LS version. How to select a proper value of  $\alpha$  is out of the scope of this paper, but the reader may refer to [23], [24].

#### B. Weighted Constrained Least Squares (WCLS)

The weighted constrained least squares (WCLS) approach employs a modification of the metric structures in the observation  $Y$  and signal  $S$  spaces, induced by the weighted square norms [22]

$$\|\mathbf{y} - \mathbf{A}\mathbf{s}\|_Y^2 = [(\mathbf{y} - \mathbf{A}\mathbf{s}), (\mathbf{y} - \mathbf{A}\mathbf{s})]_Y = (\mathbf{y} - \mathbf{A}\mathbf{s})^+ \mathbf{W}_Y (\mathbf{y} - \mathbf{A}\mathbf{s}), \quad (16)$$

$$\|\mathbf{s}\|_S^2 = [\mathbf{s}, \mathbf{s}]_S = \mathbf{s}^+ \mathbf{W}_S \mathbf{s}. \quad (17)$$

where  $\mathbf{W}_Y$  and  $\mathbf{W}_S$  are symmetric definite-positive invertible (well-conditioned) weighting matrices.

Prior knowledge about the form of the solution is embedded in the weighting matrix  $\mathbf{W}_Y$ , which makes the selection of such matrix problem dependent.

Following the same strategy as the CLS approach, the WCLS optimization problem is defined by

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \left\{ \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_Y^2 + \alpha \|\mathbf{s}\|_S^2 \right\}. \quad (18)$$

Here, the modified augmented objective function  $\Lambda(\mathbf{s})$  is given by

$$\Lambda(\mathbf{s}) = \mathbf{s}^+ \mathbf{A}^+ \mathbf{W}_Y \mathbf{A} \mathbf{s} - \mathbf{y}^+ \mathbf{W}_Y \mathbf{A} \mathbf{s} - \mathbf{s}^+ \mathbf{A}^+ \mathbf{W}_Y \mathbf{y} + \alpha (\mathbf{s}^+ \mathbf{W}_S \mathbf{s}), \quad (19)$$

where the terms independent to vector  $\mathbf{s}$  have been ignored. Applying the standard gradient method to solve (18) with respect to  $\mathbf{s}$ , results in the derivative expression,

$$\frac{\partial}{\partial \mathbf{s}} \Lambda(\mathbf{s}) = -\mathbf{A}^+ \mathbf{W}_Y (\mathbf{y} - \mathbf{A}\mathbf{s}) + \alpha \mathbf{W}_S \mathbf{s} = \mathbf{0}, \quad (20)$$

that yields the variational equation

$$(\mathbf{A}^+ \mathbf{W}_Y \mathbf{A} + \alpha \mathbf{W}_S) \mathbf{s} = \mathbf{A}^+ \mathbf{W}_Y \mathbf{y}. \quad (21)$$

The WCLS solver  $\hat{\mathbf{s}} = \mathbf{F}_{\text{WCLS}} \mathbf{y}$  has the same form as the CLS estimator, but with a different solution operator

$$\mathbf{F}_{\text{WCLS}} = (\mathbf{A}^+ \mathbf{W}_Y \mathbf{A} + \alpha \mathbf{W}_S)^{-1} \mathbf{A}^+ \mathbf{W}_Y, \quad (22)$$

which can also be used in (8) to retrieve an estimate  $\hat{\mathbf{b}}_{\text{WCLS}}$ .

The WCLS algorithm offers additional degrees-of-freedom at the signal processing level. These are determined by the weights that induce the metrics in the observation  $Y$  and signal  $S$  spaces. The setting of different weights ( $\mathbf{W}_Y \neq \mathbf{I}$ ,  $\mathbf{W}_S \neq \mathbf{I}$ ) imposes prior knowledge on the solution via regularization, making the desired solution smoother and continuously dependent on the data.

#### IV. STATISTICAL REGULARIZATION

In the statistical regularization methodology, the inference is made in terms of probabilistic statements, where the Bayes minimum risk (BMR) estimation strategy plays a key role [17] – [19], [25].

##### A. Bayes Minimum Risk (BMR)

The main distinction between the Bayes strategy and the presented descriptive regularization methodology is in the use of probabilistic models that quantify the uncertainty of the unknowns. The signal  $\mathbf{s}$  and the observed data  $\mathbf{y}$  are intrinsically random processes, characterized by their probability distributions.

We first introduce the loss function [22], [24],

$$f(\mathbf{s}, \mathbf{F}\mathbf{y}) = \|\mathbf{F}\mathbf{y} - \mathbf{s}\|^2, \quad (23)$$

that represents a measure of loss or discrepancy between the desired information (signal) vector  $\mathbf{s}$  and its actual estimate

$\hat{\mathbf{s}} = \mathbf{F}\mathbf{y}$ . The Bayes risk functional is the expected value of the loss  $r(\mathbf{F}) = \mathbb{E}(\|\mathbf{F}(\mathbf{A}\mathbf{s} + \mathbf{n}) - \mathbf{s}\|^2)$ .

The BMR strategy is formulated as that of deriving the linear operator  $\mathbf{F}$  that provides a solution to the following optimization problem [17], [18], [25],

$$\mathbf{F} = \arg \min_{\mathbf{F}} \left\{ \mathbb{E}(\|\mathbf{F}(\mathbf{A}\mathbf{s} + \mathbf{n}) - \mathbf{s}\|^2) \right\}. \quad (24)$$

Notice that  $r(\mathbf{F})$  can be decomposed in terms of inner products as

$$r(\mathbf{F}) = \mathbb{E}(\mathbf{s}^+ (\mathbf{F}\mathbf{A} - \mathbf{I})^+ (\mathbf{F}\mathbf{A} - \mathbf{I}) \mathbf{s}) + \mathbb{E}(\mathbf{n}^+ \mathbf{F}^+ \mathbf{F} \mathbf{n}). \quad (25)$$

Using the following properties [17], [18],

$$\mathbf{x}^+ \mathbf{Q} \mathbf{x} = \text{tr} \{ \mathbf{Q} \mathbf{x} \mathbf{x}^+ \}, \quad (26)$$

$$\text{tr} \{ \mathbf{O} \mathbf{P} \mathbf{Q} \} = \text{tr} \{ \mathbf{Q} \mathbf{O} \mathbf{P} \} = \text{tr} \{ \mathbf{P} \mathbf{Q} \mathbf{O} \}, \quad (27)$$

in which  $\mathbf{x}$  is an arbitrary vector and  $\mathbf{O} \mathbf{P} \mathbf{Q}$  are square matrices, yields

$$r(\mathbf{F}) = \underbrace{\text{tr} \{ (\mathbf{F}\mathbf{A} - \mathbf{I})^+ \mathbb{E}(\mathbf{s}^+ \mathbf{s}) (\mathbf{F}\mathbf{A} - \mathbf{I}) \}}_{r_s(\mathbf{F})} + \underbrace{\text{tr} \{ \mathbf{F}^+ \mathbb{E}(\mathbf{n}^+ \mathbf{n}) \mathbf{F} \}}_{r_n(\mathbf{F})}. \quad (28)$$

Here, the first term at right hand  $r_s(\mathbf{F})$  represents the measure of discrepancy between the composition  $\mathbf{F}\mathbf{A}$  and  $\mathbf{I}$ , meaning how the solution operator  $\mathbf{F}$  differs from the system pseudoinverse. While  $r_n(\mathbf{F})$  is the measure of noise energy in the resultant solution.

To determine the optimum operator  $\mathbf{F}$ , we differentiate the risk functional  $r(\mathbf{F})$  with respect to  $\mathbf{F}$  (see the Appendix) and set the result to zero,

$$\frac{\partial}{\partial \mathbf{F}} r(\mathbf{F}) = 2(\mathbf{F}\mathbf{A} - \mathbf{I}) \mathbb{E}(\mathbf{s}^+ \mathbf{s}) \mathbf{A}^+ + 2\mathbf{F} \mathbb{E}(\mathbf{n}^+ \mathbf{n}) = \mathbf{0}. \quad (29)$$

Rearranging the terms in (29),

$$\mathbf{F} (\mathbf{A} \mathbb{E}(\mathbf{s}^+ \mathbf{s}) \mathbf{A}^+ + \mathbb{E}(\mathbf{n}^+ \mathbf{n})) = \mathbb{E}(\mathbf{s}^+ \mathbf{s}) \mathbf{A}^+, \quad (30)$$

and recognizing that  $\mathbb{E}(\mathbf{s}^+ \mathbf{s}) = \mathbf{R}_s$  in (4) and  $\mathbb{E}(\mathbf{n}^+ \mathbf{n}) = \mathbf{R}_n$  in (5), we derive the BMR optimal (in a statistical sense) solution operator,

$$\mathbf{F}_{\text{BMR}} = \mathbf{R}_s \mathbf{A}^+ (\mathbf{A} \mathbf{R}_s \mathbf{A}^+ + \mathbf{R}_n)^{-1}, \quad (31)$$

used in (8) to retrieve an estimate  $\hat{\mathbf{b}}_{\text{BMR}}$ .

The BMR methodology results in the tractable solution for any fixed but unknown probabilities of the desired vector and data.

The main disadvantage of such approach relies on the a priori knowledge of the correlation matrices  $\mathbf{R}_s$  and  $\mathbf{R}_n$ . Factor  $N_0$  in (5) is typically evaluated directly from the actually acquired measurement data  $\mathbf{y}$ . While,  $\mathbf{R}_s = \mathbf{D}(\hat{\mathbf{b}})$  in (4) depends on a first estimate of the PSP  $\hat{\mathbf{b}}^{[0]}$ , which makes  $\mathbf{F}_{\text{BMR}}$  a solution-dependent solution operator.

The BMR strategy provides the overall optimality of the solution, which is the minimum value of the Bayesian risk. The latter refers to the compromise between the systematic and fluctuation errors in the resulting solution.

### B. Maximum-likelihood (ML)

Due to the intrinsic statistical nature of the physical phenomena, responsible of forming the signal  $\mathbf{s}$ , noise  $\mathbf{n}$  and data  $\mathbf{y}$  vectors, it is customary to model them as random (stochastic) vectors.

A statistical vector represents, in theory, an infinite number of different realizations of a process. Such vector is explicitly characterized through its *probability density function* (pdf). The pdf of the  $L$ -dimensional Gaussian zero-mean complex-valued vector  $\mathbf{y}$  is given by [12], [13], [26],

$$p(\mathbf{y}) = \pi^{-L} \det^{-1}\{\mathbf{R}_y\} \exp\left\{-\left(\mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y}\right)\right\}. \quad (32)$$

In order to retrieve a statistically optimal solution to the previously specified non-linear ill-posed TomoSAR inverse problem, the addressed BMR strategy is extended to the maximum a posteriori (MAP) probability approach [22], by imposing no constraints on linearity and by assuming a priori Gaussian statistics of the noise  $\mathbf{n}$  and PSP  $\mathbf{b}$  vectors, that results in

$$\hat{\mathbf{b}}_{\text{MAP}} = \arg \max_{\mathbf{b}} \{p(\mathbf{b} | \mathbf{y})\}. \quad (33)$$

Using the pdf Bayes formula, we express the conditional pdf as

$$p(\mathbf{b} | \mathbf{y}) = (1/p(\mathbf{y})) p(\mathbf{y} | \mathbf{b}) p(\mathbf{b}). \quad (34)$$

While its logarithm is given by

$$\ln p(\mathbf{b} | \mathbf{y}) = \ln p(\mathbf{y} | \mathbf{b}) + \ln p(\mathbf{b}) - \ln p(\mathbf{y}). \quad (35)$$

In scenarios where the a priori pdf  $p(\mathbf{b})$  is unknown (which is the case under study) and set to a constant distribution, the MAP strategy reduces to the ML methodology. Thus, setting  $p(\mathbf{b}) \approx \text{const}$  and ignoring the

terms that do not comprise  $\mathbf{b}$  in (35), we define the log-likelihood function of vector  $\mathbf{b}$  as the logarithm of the conditional pdf  $p(\mathbf{y} | \mathbf{b})$

$$\ln p(\mathbf{y} | \mathbf{b}) = -\ln \det\{\mathbf{R}_y\} - \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y}, \quad (36)$$

where the terms that do not contain  $\mathbf{b}$  have been ignored. The ML solution is then reduced to the minimization problem

$$\hat{\mathbf{b}}_{\text{ML}} = \arg \min_{\mathbf{b}} \{\Lambda(\mathbf{b})\}, \quad (37)$$

with the objective function  $\Lambda(\mathbf{b}) = -\ln p(\mathbf{y} | \mathbf{b})$ .

In [13], equation (37) is solved using the standard gradient method, yielding a technique named MARIA (ML-inspired adaptive robust iterative approach). This paper solves (37) in a different manner, whose new developments play an important role in the succeeding derivation of WISE.

First, as in [13], we apply the standard gradient method to solve (37) with respect to  $\mathbf{b}$  (refer to the Appendix), resulting in the derivative expressions,

$$\frac{\partial}{\partial \mathbf{b}} \ln \det\{\mathbf{R}_y\} = \{\mathbf{A}^+ \mathbf{R}_y^{-1} \mathbf{A}\}_{\text{diag}}, \quad (38)$$

$$\frac{\partial}{\partial \mathbf{b}} \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y} = -\{\mathbf{A}^+ \mathbf{R}_y^{-1} \mathbf{y} \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{A}\}_{\text{diag}}. \quad (39)$$

Recall that  $\mathbf{R}_y = \mathbf{R}_y(\mathbf{b}) = \mathbf{A} \mathbf{D}(\mathbf{b}) \mathbf{A}^+ + \mathbf{R}_n$ .

Consider the next property for a diagonal matrix  $\mathbf{D}(\mathbf{b})$  and an arbitrary square matrix  $\mathbf{Q}$  [13], [17], [19],

$$\mathbf{D}(\mathbf{b}) \{\mathbf{Q}\}_{\text{diag}} = \{\mathbf{D}(\mathbf{b}) \mathbf{Q}\}_{\text{diag}}. \quad (40)$$

Setting the derivative of  $\Lambda(\mathbf{b})$  to zero, rearranging and multiplying both sides by  $\mathbf{D}(\mathbf{b})^2$  yields to

$$\{\mathbf{D}(\mathbf{b}) \mathbf{A}^+ \mathbf{R}_y^{-1} \mathbf{y} \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{A} \mathbf{D}(\mathbf{b})\}_{\text{diag}} = \{\mathbf{D}(\mathbf{b}) \mathbf{A}^+ \mathbf{R}_y^{-1} \mathbf{A} \mathbf{D}(\mathbf{b})\}_{\text{diag}}. \quad (41)$$

Now, taking into account the Capon spectral estimate [3], [8], [9], of the PSP vector  $\mathbf{b}$ ,

$$\left\{ \hat{b}_{\text{Capon}_m} = \frac{1}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \right\}_{m=1}^M, \quad (42)$$

and performing a sequence of evident manipulations, equation (41) can also be expressed as follows

$$\left\{ \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{y} \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{a}_m}{(\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m)^2} = \frac{1}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \right\}_{m=1}^M, \quad (43)$$

which holds for  $\mathbf{R}_y = \mathbf{Y} = \mathbf{y} \mathbf{y}^+$ .

Since  $\mathbf{b} \approx \hat{\mathbf{b}}_{\text{Capon}}$ , an approximate solution to (37) can be retrieved. The latter leads to the formulation of the ML-based method [26, Eq. 32],

$$\left\{ \hat{\mathbf{b}}_{\text{ML}_m} = \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{(\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m)^2} \right\}_{m=1}^M, \quad (44)$$

with  $\mathbf{Y} = \mathbf{y} \mathbf{y}^+$ .

Equation (44) can also be represented as a non-linear matrix-vector equation  $\hat{\mathbf{b}}_{\text{ML}} = \{\mathbf{F}_{\text{ML}} \mathbf{Y} \mathbf{F}_{\text{ML}}^+\}_{\text{diag}}$  with

$$\mathbf{F}_{\text{ML}} = \mathbf{D}(\hat{\mathbf{b}}) \mathbf{A}^+ \mathbf{R}_y^{-1}, \quad (45)$$

and  $\hat{\mathbf{b}}$ , at the principal diagonal of  $\mathbf{D}(\hat{\mathbf{b}})$ , recovered using (42). It can be observed that, the corresponding solution operator  $\mathbf{F}_{\text{ML}}$  is also solution-dependent, as the one utilized by the BMR solver in (31).

In the MAP estimation strategy, we involve a priori information about the desired PSP vector  $\mathbf{b}$  in the form of its a priori pdf  $p(\mathbf{b})$ . This leads to the augmented criterion functional in (35), where the second term pertains to our prior knowledge.

In contrast, in the proposed ML estimation strategy, we do not impose any a priori knowledge about the statistical distribution of the desired PSP vector  $\mathbf{b}$ . Information about vector  $\mathbf{b}$  is extracted only from the measurement data, as expressed in (36).

Equation (44) is the basis for several approaches; differ one from each other depending on the assumptions and considerations made during the respective implementations.

The solution to the Capon spatial filter design problem [8]

$$\left\{ \arg \min_{\mathbf{h}_m} \{ \mathbf{h}_m^+ \mathbf{R}_y \mathbf{h}_m \} \text{ s.t. } \mathbf{h}_m^+ \mathbf{a}_m = 1 \right\}_{m=1}^M, \quad (46)$$

with  $\{\mathbf{h}_m\}_{m=1}^M$  as weighting vectors, is actually given by (44). The classical Capon beamforming [8, Eq. 6.3.26] simply assumes that  $\mathbf{R}_y = \mathbf{Y} = \mathbf{y} \mathbf{y}^+$ .

Derived differently, equation (44) is also the basis for the iterative adaptive approach for ML (IAA-ML) introduced in [26, Eq. 32]. The IAA-ML works well with arbitrary array geometries and can be extended to give sparse results via model-order selection [26].

The wavelet domain thresholding (WDT)-refined virtual adaptive beamforming (WAVAB) technique introduced in [27, Eq. 19 and Eq. 20] solves (44) in an adaptive fixed-point iterative fashion that do not involve any matrix inversion at all

iteration steps, which differs from Capon beamforming that requires inversions of  $\mathbf{Y}$  and from IAA-ML that requires inversions of  $\mathbf{R}_y$ .

Finally, the implementation of (44) introduced afterward in Section VI is presented as a different version of MARIA [13], since, as in the previous version, it refers to an adaptive iterative ML-based approach, which refines, after each iteration, a first estimate of the PSP. It performs resolution-enhancement, suppression of artifacts and ambiguity levels reduction. Treated as an extension of Capon (refer to (43)), we suggest recovering the first PSP estimate using Capon beamforming, however, as we will observe later, the algorithm allows for the usage of any TomoSAR focusing technique for this purpose.

### C. Weighted covariance fitting (WCF) criterion

If we relax the Gaussian assumption in (32), then the estimates given after solving (37) are no longer ML estimates. In such a case, it makes more sense to use a covariance fitting criterion. Thus, taking into consideration the developments of the previously addressed regularization approaches, a new method based on the WCF criterion [28] – [30] is introduced.

For the single-look case, the WCF criterion refers to the following optimization problem [28] – [30],

$$\hat{\mathbf{b}}_{\text{WCF}} = \arg \min_{\mathbf{b}} \left\{ \left\| \mathbf{R}_y^{-1/2} (\mathbf{y} \mathbf{y}^+ - \mathbf{R}_y) \right\|^2 \right\}. \quad (47)$$

Recall that  $\mathbf{R}_y = \mathbf{R}_y(\mathbf{b}) = \mathbf{A} \mathbf{D}(\mathbf{b}) \mathbf{A}^+ + \mathbf{R}_n$ .

The expansion of the objective function  $\Lambda(\mathbf{b})$  is given by [29],

$$\Lambda(\mathbf{b}) = \|\mathbf{y}\|^2 \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y} + \text{tr} \{ \mathbf{R}_y \}, \quad (48)$$

where the terms independent to vector  $\mathbf{b}$  have been ignored.

It follows from (48), that the minimization of (47) is equivalent to the minimization of

$$\Lambda(\mathbf{b}) = \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y} + \frac{\text{tr} \{ \mathbf{R}_y \}}{\text{tr} \{ \mathbf{Y} \}}, \quad (49)$$

since

$$\|\mathbf{y}\|^2 = [\mathbf{y}, \mathbf{y}] = \text{tr} \{ \mathbf{y} \mathbf{y}^+ \}, \quad (50)$$

and  $\mathbf{Y} = \mathbf{y} \mathbf{y}^+$ , as previously defined in (9).

In [28], [29], (47) is solved with the aim of introducing the semiparametric/sparse iterative covariance-based estimation (SPICE) method. Yet, instead of deriving a line-spectrum estimator as SPICE, this paper presents a novel non-parametric approach.

Keeping in mind the developments that yield equation (44), from which the non-parametric approaches Capon, IAA-ML

and WAVAB are inferred, this paper provides a solution to (47), which has a similar structure to (44).

We apply the standard gradient method to solve (47) with respect to  $\mathbf{b}$ . According to the earlier addressed ML approach, the derivative expression of the first term at right hand of the objective function in (49), yields

$$\frac{\partial}{\partial \mathbf{b}} \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y} = - \left\{ \mathbf{A}^+ \mathbf{R}_y^{-1} \mathbf{y} \mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{A} \right\}_{\text{diag}}. \quad (51)$$

Next, consider the derivative operation (refer to the Appendix)

$$\frac{\partial}{\partial \mathbf{b}} \text{tr} \{ \mathbf{R}_y \} = \frac{\partial}{\partial \mathbf{b}} \text{tr} \{ \mathbf{A} \mathbf{D}(\mathbf{b}) \mathbf{A}^+ \} = \mathbf{A}^+ \mathbf{A}, \quad (52)$$

where matrix  $\mathbf{A}$  is not function of vector  $\mathbf{b}$ , and the term that does not comprise  $\mathbf{b}$  is ignored. Here  $\mathbf{\Psi} = \mathbf{A}^+ \mathbf{A}$  is recognized to be the discrete-form ambiguity function matrix operator, which describes the distortions due to focusing [31].

Taking into account the derivative expressions in (51) and (52), we set the derivative of  $\Lambda(\mathbf{b})$  to  $\mathbf{0}$ , and we express it as

$$\left\{ \mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m = \frac{1}{\text{tr} \{ \mathbf{Y} \}} \mathbf{a}_m^+ \mathbf{a}_m \right\}_{m=1}^M. \quad (53)$$

As in the prior ML case, we make use of the Capon spectral estimate [3], [8], [9], of the PSP vector  $\mathbf{b}$  in (42) to derive a WCF-based solver. Multiplying both sides in (53) by (42), and performing a sequence of evident manipulations, gives

$$\left\{ \frac{\text{tr} \{ \mathbf{Y} \}}{\mathbf{a}_m^+ \mathbf{a}_m} \left( \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \right) = \frac{1}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \right\}_{m=1}^M. \quad (54)$$

The term at right hand of equation (54) is recognized to be the Capon's spectral estimate of the PSP,  $\mathbf{b} \approx \hat{\mathbf{b}}_{\text{Capon}}$ , leading to an approximate solution to (47). The novel proposed WCF-inspired solver is then given by

$$\left\{ \hat{\mathbf{b}}_{\text{WCF}_m} = \frac{\text{tr} \{ \mathbf{Y} \}}{\mathbf{a}_m^+ \mathbf{a}_m} \left( \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \right) \right\}_{m=0}^M. \quad (55)$$

The introduced approach results to be also solution dependent, since, by definition, matrix  $\mathbf{R}_y$  in (6) depends on a first estimate of the PSP  $\hat{\mathbf{b}}^{[0]}$ . As in (43), the term at right hand of equation (54) represents the asymptotic of the filter output given on the left hand side.

## V. SIMILARITY BETWEEN THE ADDRESSED REGULARIZATION APPROACHES

The WCF-inspired method in (55) has been derived as a consequence to the developments employed to yield the

related addressed regularization techniques. This section treats the relationship between the aforementioned deterministic and statistical estimators.

In order to reveal the similarity between the addressed techniques, we derive a second representation of the optimal BMR solution operator  $\mathbf{F}_{\text{BMR}}$  in (31). Employing the Frobenius formula for inverting the matrix of special form [13], [17], [18],

$$\mathbf{R}_y^{-1} = \mathbf{R}_n^{-1} - \mathbf{R}_n^{-1} \mathbf{A} \mathbf{U} \mathbf{A}^+ \mathbf{R}_n^{-1}, \quad (56)$$

where

$$\mathbf{U} = \left( \mathbf{A}^+ \mathbf{R}_n^{-1} \mathbf{A} + \mathbf{R}_s^{-1} \right)^{-1}; \quad (57)$$

and performing some mathematical manipulations, we get

$$\begin{aligned} \mathbf{F}_{\text{BMR}} &= \mathbf{R}_s \mathbf{A}^+ \left( \mathbf{R}_n^{-1} - \mathbf{R}_n^{-1} \mathbf{A} \mathbf{U} \mathbf{A}^+ \mathbf{R}_n^{-1} \right) \\ &= \mathbf{R}_s \left( \mathbf{U}^{-1} - \mathbf{A}^+ \mathbf{R}_n^{-1} \mathbf{A} \right) \mathbf{U} \mathbf{A}^+ \mathbf{R}_n^{-1} \\ &= \mathbf{U} \mathbf{A}^+ \mathbf{R}_n^{-1}. \end{aligned} \quad (58)$$

Consider now the  $\delta$ -correlation simplest model of noise and signal,  $\mathbf{R}_n = N_0 \mathbf{I}$ ,  $\mathbf{R}_s = S_0 \mathbf{I}$ , in which case

$$\mathbf{F}_{\text{BMR}} = \left( \mathbf{A}^+ \mathbf{A} + \frac{N_0}{S_0} \mathbf{I} \right)^{-1} \mathbf{A}^+, \quad (59)$$

has the same structure of the CLS solution operator in (15), but with a regularization parameter  $\alpha = N_0 / S_0$ , which provides a balance between the systematic and noise error measures.

On the other hand, any of both WCLS and BMR methods can be referred to as the particular case of another under some model assumptions. Putting  $\alpha \mathbf{W}_s = \mathbf{R}_s^{-1}$  and  $\mathbf{W}_y = \mathbf{R}_n^{-1}$  yields

$$\mathbf{F}_{\text{WCLS}} = \mathbf{F}_{\text{BMR}} = \left( \mathbf{A}^+ \mathbf{R}_n^{-1} \mathbf{A} + \mathbf{R}_s^{-1} \right)^{-1} \mathbf{A}^+ \mathbf{R}_n^{-1}. \quad (60)$$

In this way, the deterministic WCLS technique provides a statistically optimal solution, in the BMR sense, to the TomoSAR inverse problem. When the prior knowledge of  $\mathbf{R}_n$  and  $\mathbf{R}_s$  is available, then the WCLS method provides an estimator which performs the statistically best behavior. In other cases, the WCLS method provides the robust approximation of the statistically optimal BMR technique. In such instance, the unknown models  $\mathbf{R}_n$  and  $\mathbf{R}_s$ , can be approximated by adjusting the degrees-of-freedom,  $\mathbf{W}_y$  and  $\alpha \mathbf{W}_s$ .

The addressed ML approach is treated as an extension of the BMR strategy. This is noticeable, since both ML and BMR solution operators  $\mathbf{F}$ , have the same structure as in (31) (refer to the definition of  $\mathbf{R}_s$  in (4) and  $\mathbf{R}_y$  in (6)). The difference



TABLE I  
SIMILARITY BETWEEN THE ADDRESSED REGULARIZATION APPROACHES

Optimization Criterion	Regularization Technique		Particular Cases	Condition
CLS*	$\hat{\mathbf{b}} = \left\{ \mathbf{F} \mathbf{Y} \mathbf{F}^+ \right\}_{\text{diag}}$	$\mathbf{F}_{\text{CLS}} = \left( \mathbf{A}^+ \mathbf{A} + \alpha \mathbf{I} \right)^{-1} \mathbf{A}^+$	$\mathbf{F}_{\text{CLS}} = \mathbf{F}_{\text{LS}}$	$\alpha \rightarrow 0$
WCLS*		$\mathbf{F}_{\text{WCLS}} = \left( \mathbf{A}^+ \mathbf{W}_S \mathbf{A} + \alpha \mathbf{W}_Y \right)^{-1} \mathbf{A}^+ \mathbf{W}_S$	$\mathbf{F}_{\text{WCLS}} = \mathbf{F}_{\text{CLS}}$	$\mathbf{W}_S = \mathbf{I}$ and $\mathbf{W}_Y = \mathbf{I}$
BMR**		$\mathbf{F}_{\text{BMR}} = \left( \mathbf{A}^+ \mathbf{R}_n^{-1} \mathbf{A} + \mathbf{R}_s^{-1} \right)^{-1} \mathbf{A}^+ \mathbf{R}_n^{-1}$	$\mathbf{F}_{\text{BMR}} = \mathbf{F}_{\text{CLS}}$	$\mathbf{R}_n = N_0 \mathbf{I}$ and $\mathbf{R}_s = S_0 \mathbf{I}$ ; $\alpha = N_0 / S_0$
			$\mathbf{F}_{\text{BMR}} = \mathbf{F}_{\text{WCLS}}$	$\mathbf{R}_s^{-1} = \alpha \mathbf{W}_S$ and $\mathbf{R}_n^{-1} = \mathbf{W}_Y$
		$\mathbf{F}_{\text{BMR}} = \mathbf{R}_s \mathbf{A}^+ \left( \mathbf{A} \mathbf{R}_s \mathbf{A}^+ + \mathbf{R}_n \right)^{-1}$	$\mathbf{F}_{\text{BMR}} = \mathbf{F}_{\text{ML}}$	$\hat{\mathbf{b}} = \left\{ \hat{b}_{\text{Capon}_m} \right\}_{m=1}^M$ in $\mathbf{R}_s = \mathbf{D} \left( \hat{\mathbf{b}} \right)$
ML**	$\left\{ \hat{b}_{\text{ML}_m} = \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\left( \mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m \right)^2} \right\}_{m=1}^M$		$\hat{\mathbf{b}}_{\text{ML}} = \hat{\mathbf{b}}_{\text{Capon}}$	$\mathbf{R}_y = \mathbf{Y}$
WCF**	$\left\{ \hat{b}_{\text{WCF}_m} = \frac{\text{tr} \left\{ \mathbf{Y} \right\}}{\mathbf{a}_m^+ \mathbf{a}_m} \left( \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \right) \right\}_{m=0}^M$		$\hat{\mathbf{b}}_{\text{WCF}} = \hat{\mathbf{b}}_{\text{Capon}}$	$\left\{ \text{tr} \left\{ \mathbf{Y} \right\} \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{a}_m} = 1 \right\}_{m=0}^M$

\* Descriptive regularization \*\* Statistical regularization

lies in the content of the principal diagonal of the diagonal matrices  $\mathbf{D}(\hat{\mathbf{b}})$  in (31) and (45) respectively. For the ML case, the main diagonal of matrix  $\mathbf{D}(\hat{\mathbf{b}})$  in (45) is composed of a first estimate of the PSP  $\hat{\mathbf{b}}^{[0]}$ , retrieved using the Capon approach given by equation (42). While, for the BMR case, there are no constraints in the choice of the focusing technique in charge of recovering a first estimate  $\hat{\mathbf{b}}^{[0]}$ . Therefore, the ML approach can be stated as a particular case of the BMR method.

Notice that, the ML-inspired solver in (44) becomes the Capon method in (42) when  $\mathbf{Y} = \mathbf{R}_y$ . Meaning, that the addressed ML-based technique can be also seen as a variant (improvement) of the Capon approach, adapted to the case when there are discrepancies between the measured data covariance matrix  $\mathbf{Y}$  and the actual (modelled) data covariance matrix  $\mathbf{R}_y$ .

The WCF minimization problem in (47) relates to the addressed ML optimization problem in (37), in that both objective functions share the same term  $\mathbf{y}^+ \mathbf{R}_y^{-1} \mathbf{y}$ . Since both minimization problems are tackled applying the standard gradient method, the resultant derivative expression of such term results to be the same.

Following the developments that derived into the ML technique, the WCF solver in (55) is presented as well as a modification to the Capon approach in (42), but employing the WCF criterion. For the particular case, when  $\mathbf{Y} = \mathbf{R}_y$ , an estimate of the PSP  $\hat{\mathbf{b}}$  is recovered using only the information extracted from the measured covariance matrix  $\mathbf{Y}$  and the ambiguity function matrix operator  $\Psi$ .

Finally, for an easy assessment, a synthesized view of the addressed regularization approaches is depicted in Table I.

## VI. ITERATIVE ADAPTIVE IMPLEMENTATION

In the previous related studies [12], [13], the TomoSAR inverse problem is tackled using statistical regularization approaches based on ML. Being (44) a ML-inspired technique, which also presents a similar structure to the introduced WCF-based method in (55), we choose this regularization technique for further comparisons.

This section presents the implementation details of both MARIA (inferred from (44)) and WISE (inferred from (55)) iterative adaptive approaches.

Before introducing the corresponding implementations, we perform a subtle but important modification to the respective basic equations, given by

$$\left\{ \hat{b}_{\text{MARIA}_m} = \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \hat{b}_m \right\}_{m=1}^M, \quad (60)$$

and

$$\left\{ \hat{b}_{\text{WISE}_m} = \text{tr}\{\mathbf{Y}\} \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{a}_m} \hat{b}_m \right\}_{m=0}^M, \quad (61)$$

respectively.

The introduced WISE technique, along with MARIA, is a statistical regularization approach that requires of the PSP vector  $\mathbf{b}$  as a priori information. The iterative adaptive implementations of such methods, given in Table II and Table III, correspondingly, are considered as sub-optimal versions of the ML technique in (44) and the WCF method in (55), since the actual PSP vector  $\mathbf{b}$  is unknown and needs to be firstly estimated ( $\hat{\mathbf{b}}^{[0]} \neq \mathbf{b}$ ).

Because of consistency reasons, the first estimate  $\hat{\mathbf{b}}^{[0]}$  is retrieved using the conventional Capon beamforming technique, which in a matrix-vector form is defined by [19]

$$\hat{\mathbf{b}}_{\text{Capon}} = \left\{ \left[ \mathbf{D} \left( \left\{ \mathbf{A}^+ \mathbf{Y}^{-1} \mathbf{A} \right\}_{\text{diag}} \right) \right]^{-1} \right\}_{\text{diag}} ; \quad (62)$$

notice that, in contrast to (42), here we assume that  $\mathbf{R}_y = \mathbf{Y}$ . Consequently, the degradation or improvement of the retrieved solutions depends on the precision of the first estimate  $\hat{\mathbf{b}}_{\text{Capon}} = \hat{\mathbf{b}}^{[0]}$ . The latter entails the constraint that a high number of looks is required in order to avoid the rank-deficiencies of the data covariance matrix  $\mathbf{Y}$  and to assure the best performance of Capon.

Another drawback is that the dependence on a first estimate  $\hat{\mathbf{b}}^{[0]}$  implies that no unique regularization method to recover  $\hat{\mathbf{b}}$  in (60) and (61) exists [13], [18]. Different solutions to the TomoSAR non-linear inverse problem are retrieved for different estimates  $\hat{\mathbf{b}}^{[0]}$ , i.e., when it is in some way corrupted. To cope with this problem, such MARIA and WISE estimators are implemented in an adaptive (with respect to the desired solution) iterative manner, as shown in Tables II and III. Nonetheless, this only alleviates the problem until certain extent, when the discrepancies between the different estimates  $\hat{\mathbf{b}}^{[0]}$  are not highly significant.

At first instance, notice that, in comparison to IAA-ML [26, Table II] the MARIA (Table II) and WISE (Table III) algorithms have a different initialization. The IAA-ML is restricted to the standard delay-and-sum beamformer, whereas MARIA and WISE utilize Capon beamforming. Nonetheless, the algorithms allow for the usage of any TomoSAR focusing technique for the retrieval of the first estimate of the PSP.

The main difference between IAA-ML and the MARIA and WISE algorithms is related to the modified expressions in (60) and (61), respectively. The first entry to IAA-ML is strictly employed to construct the data correlation matrix  $\mathbf{R}_y$ , which is used later to re-compute a new PSP estimate through (44). This process is repeated on each iteration. In contrast, for the cases of MARIA and WISE, the first-input (zero-step iteration) PSP estimate is not re-calculated but rather refined after each iteration. Furthermore, in order to guarantee convergence, we incorporate the projector onto convex solution sets (POCS) operator.

Since the minimization problems to be solved are convex, it is expected that they converge under the conditions that  $\{b_m^{[i]} \geq 0\}_{m=1}^M$  and that  $\mathbf{R}_y$  remains positive definite, as the number of iterations increases [28]. Hence, in order to guarantee convergence, we incorporate the POCS operator  $\mathcal{P}_{+\gamma}\{\cdot\}$  into the iterative solvers, that clips off all entries lower than the user specified nonnegative tolerance threshold,  $\gamma \geq 0$ , in the solution space [21],

TABLE II  
THE ML-INSPIRED ITERATIVE ADAPTIVE APPROACH (MARIA)

---

```

i = 0
 $\hat{\mathbf{b}}^{[i]} = \hat{\mathbf{b}}_{\text{Capon}} = \left\{ \left[ \mathbf{D} \left( \left\{ \mathbf{A}^+ \mathbf{Y}^{-1} \mathbf{A} \right\}_{\text{diag}} \right) \right]^{-1} \right\}_{\text{diag}}$ 
do
   $\mathbf{R}_y^{-1} = \left( \mathbf{A} \mathbf{D} \left( \hat{\mathbf{b}}^{[i]} \right) \mathbf{A}^+ + N_0 \mathbf{I} \right)^{-1}$ 
  for m = 1, ..., M
     $\hat{b}_m^{[i+1]} = \mathcal{P}_{+\gamma} \left\{ \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{a}_m} \hat{b}_m^{[i]} \right\}$ 
  i ++
while  $\left\| \hat{\mathbf{b}}^{[i+1]} - \hat{\mathbf{b}}^{[i]} \right\| > \varepsilon_{TL}$  or i ≠ iMAX
 $\hat{\mathbf{b}}_{\text{MARIA}} = \hat{\mathbf{b}}^{[i=I]}$ 

```

---

TABLE III  
THE WCF-BASED ITERATIVE SPECTRAL ESTIMATOR (WISE)

---

```

i = 0
 $\hat{\mathbf{b}}^{[i]} = \hat{\mathbf{b}}_{\text{Capon}} = \left\{ \left[ \mathbf{D} \left( \left\{ \mathbf{A}^+ \mathbf{Y}^{-1} \mathbf{A} \right\}_{\text{diag}} \right) \right]^{-1} \right\}_{\text{diag}}$ 
do
   $\mathbf{R}_y^{-1} = \left( \mathbf{A} \mathbf{D} \left( \hat{\mathbf{b}}^{[i]} \right) \mathbf{A}^+ + N_0 \mathbf{I} \right)^{-1}$ 
  for m = 1, ..., M
     $\hat{b}_m^{[i+1]} = \mathcal{P}_{+\gamma} \left\{ \text{tr} \{ \mathbf{Y} \} \frac{\mathbf{a}_m^+ \mathbf{R}_y^{-1} \mathbf{Y} \mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^+ \mathbf{a}_m} \hat{b}_m^{[i]} \right\}$ 
  i ++
while  $\left\| \hat{\mathbf{b}}^{[i+1]} - \hat{\mathbf{b}}^{[i]} \right\| > \varepsilon_{TL}$  or i ≠ iMAX
 $\hat{\mathbf{b}}_{\text{WISE}} = \hat{\mathbf{b}}^{[i=I]}$ 

```

---

$$\mathcal{P}_{+\gamma}\{x\} = \begin{cases} 0 & \text{if } x < \gamma \\ x & \text{if } x \geq \gamma \end{cases}. \quad (63)$$

Incorporation of  $\mathcal{P}_{+\gamma}\{\cdot\}$  into the implicit iterative solvers guarantees its convergence; that is a direct sequence from the fundamental theorem of POCS [21, Sect. 15.4.5]. The number of iterations needed to achieve convergence depends on the accuracy of the first input (zero-step iteration)  $\hat{\mathbf{b}}^{[0]}$ .

Factor  $N_0$  is treated as a user specified degree-of-freedom, which performs diagonal-loading to the correlation matrix  $\mathbf{R}_y$  in (6). A priori unknown reflectivity values may complicate the selection of  $N_0$  in practice. In such a case, the normalization of the data covariance matrix is recommended.

With the proposed extensions, the previous warnings about the TomoSAR inverse problem being ill-conditioned do not apply, since the whole purpose of the ML-inspired and WCF-based regularization, via iterative adaptive processing, is to cure that same ill-conditioning. As any other regularization approach, the proposed techniques retrieve only an estimate of the PSP  $\mathbf{b}$  due to the lack of enough information.

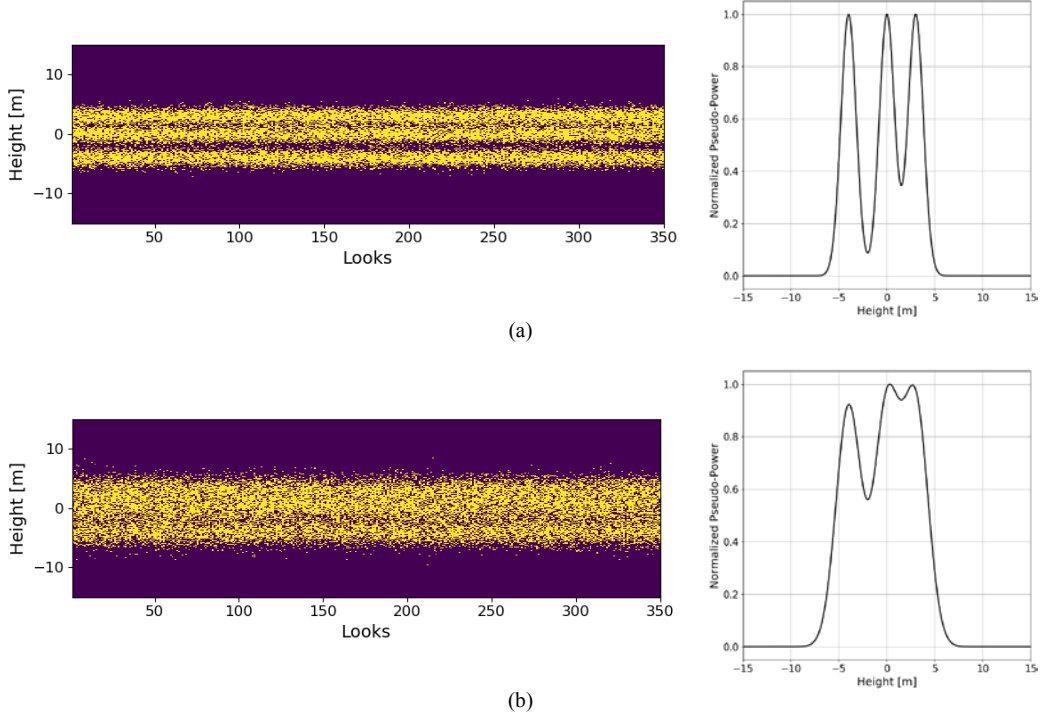


Fig. 2. Distribution of 300 scatterers displaced along the height direction, following three Gaussian distributions among  $J = 350$  independent looks, with phase-centers (means) located at  $z_1 = -4$  m,  $z_2 = 0$  m and  $z_3 = 3$  m correspondingly, and spread (standard deviation) of (a)  $\sigma = 0.8$  m and (b)  $\sigma = 1.3$  m, respectively. The theoretical distribution, for each case, is presented at the right side.

The algorithm of MARIA and the introduced WISE method, implemented in an iterative adaptive manner, as depicted in Tables II and III, respectively, is described next.

Step 1) Choose a particular value of factor  $N_0$  in  $\mathbf{R}_y$  (6), which acts as a user specified diagonal-loading parameter.

Step 2) Specify the POCS threshold value  $\gamma \geq 0$ . We recommend the simplest (zero) assignment,  $\gamma = 0$ ,  $\mathcal{P}_{+\gamma}\{\cdot\} = \mathcal{P}_{+0}\{\cdot\}$ , which defines the constraining projector onto the nonnegative convex cone solution set.

Step 3) Apply the Capon beamforming technique in (62) to get the zero-step iteration  $\hat{\mathbf{b}}^{[0]}$ . It is important to remark that a high number of looks is advisable in order to avoid the rank deficiency of the data covariance matrix  $\mathbf{Y}$  in (9).

Step 4) Compute the inversion of the diagonal-loaded (thus invertible) data correlation matrix  $\mathbf{R}_y^{-1}(\hat{\mathbf{b}}^{[i]})$  in (6), which is updated (refined) at each  $i$ -th iteration.

Step 5) Perform the preferred MARIA (60) or WISE (61) regularization approach, utilizing the previously computed matrix  $\mathbf{R}_y^{-1}(\hat{\mathbf{b}}^{[i]})$  and the previously recovered estimate  $\hat{\mathbf{b}}^{[i]}$  of the PSP vector  $\mathbf{b}$ .

Step 6) Repeat steps 4 and 5 until the respective iterative process is terminated at  $\hat{\mathbf{b}}^{[i]}$ , when the user specified  $\ell_2$ -norm convergence control tolerance level  $\varepsilon_{TL}$  is

attained at some  $i = I$ , or the maximum admissible number of iterations is reached [31].

## VII. SIMULATION RESULTS

The behavior of the novel introduced WISE tomographic estimator is first demonstrated through its application on simulated data covariance matrices  $\mathbf{Y}$  in (9), constructed from the outer product between the respective data vectors  $\mathbf{y}$  in (1) and its Hermitian conjugates, for  $J$  independent looks.

Each data covariance matrix  $\mathbf{Y}$  gathers the echoes of three hundred scatterers with equal reflectivity, displaced along the height direction. They follow (for convenience) three Gaussian distributions among independent looks, with phase-centers (means) located at  $z_1 = -4$  m,  $z_2 = 0$  m and  $z_3 = 3$  m, correspondingly. We consider a varying spread (standard deviation)  $0.1 \text{ m} \leq \sigma \leq 1.5 \text{ m}$  (Fig. 2 shows two representative cases) and a noise power spectral density  $N_0 = 0.01$ .

The simulated tomographic acquisition geometry spans a PLOS synthetic aperture of 70 m with  $L = 15$  tracks evenly distributed. The considered L-band SAR sensor (0.23 m wavelength) is assumed to be at a nominal altitude of 3000 m. We consider a slant-range distance of 5000 m between the master track and the middle target, meaning a Fourier resolution [4] of about 8.2 m.

For comparison purposes, we refer to MARIA in (60) and to the most competing commonly performed non-parametric focusing methods, which include Capon in (62) and CS [10], [11]. The CS approach presented in [10], [11] raises from the solution to the covariance-matching convex optimization

problem

$$\min_{\hat{\mathbf{b}}} \left\{ \left\| \mathbf{\Omega} \hat{\mathbf{b}} \right\|_{2,1} + \tau_1 \left\| \mathbf{A} \mathbf{D}(\hat{\mathbf{b}}) \mathbf{A}^+ - \mathbf{Y} \right\|_F^2 + \tau_2 \left\| \hat{\mathbf{b}} \right\|_{TV} \right\} \text{ s.t. } \left\{ \hat{b}_m > 0 \right\}_{m=1}^M, \quad (64)$$

where  $\mathbf{\Omega}$  is a wavelet transform sparsifying matrix. Here, the relaxation parameters  $\tau_1, \tau_2 \geq 0$ , control the trade-off between the sparsity in  $\mathbf{\Omega}$ , model mismatch and the total variation (TV) regularization term. For the presented analyses we consider the equi-balanced adjustment  $\tau_1 = \tau_2 = 0.5$ .

The aforementioned TomoSAR focusing techniques are applied to the simulated data covariance matrices  $\mathbf{Y}$  with the aim of retrieving the PSP for the respective cases. No prior model order selection is conducted. To get the approximate location of the phase-centers after tomographic focusing, we look for the three most prominent local maxima. Recall that, for a Gaussian distribution, the corresponding mean value (phase-center) is located at its maximum. Later on, we compare these results against the already known positions.

First we set  $\sigma = 0.8$ , as in Fig. 2(a), and simulate the data covariance matrix  $\mathbf{Y}$  with different number of looks in order to investigate the asymptotic properties of the Capon, CS, MARIA and WISE estimators. Fig. 3 shows the convergence properties of these focusing techniques to a lower mean square error (MSE) as the number of looks  $J$  increases, being CS the technique with the best performance. Capon presents worse results for a low number of looks, being inapplicable when matrix  $\mathbf{Y}$  is rank deficient ( $J < L$ ) due to the involved inversion of the data covariance matrix  $\mathbf{Y}$  in (62).

The MARIA and WISE techniques follow the same trend as Capon, since, by constructions, both regularization methods are presented as an extension of the Capon method, which acts as an initial input (zero-step iteration) in the corresponding adaptive iterative procedures (see Table II and Table III). The degradation or improvement of the solutions depends on the a priori information given in the form of a first estimate  $\hat{\mathbf{b}}^{[0]}$  of the PSP vector  $\mathbf{b}$ . Therefore, a high number of looks ( $J = 350$ ) is chosen, for which Capon beamforming presents enhanced performance.

Fig. 4 compares the MSE between the estimated and actual locations of the three most prominent local maxima, retrieved using the competing methods (Capon, CS and MARIA) and the novel proposed approach (WISE) for different spread values  $\sigma$ ; while Fig. 5 compares the Fréchet distance [32], [33], between the estimated and actual (theoretical) vertical profiles, correspondingly.

The Fréchet distance is commonly used to measure the similarity between two curves, taking into account the order and location of the points that compose them<sup>1</sup>.

Mathematically, the Fréchet distance between two curves is defined as follows [31],

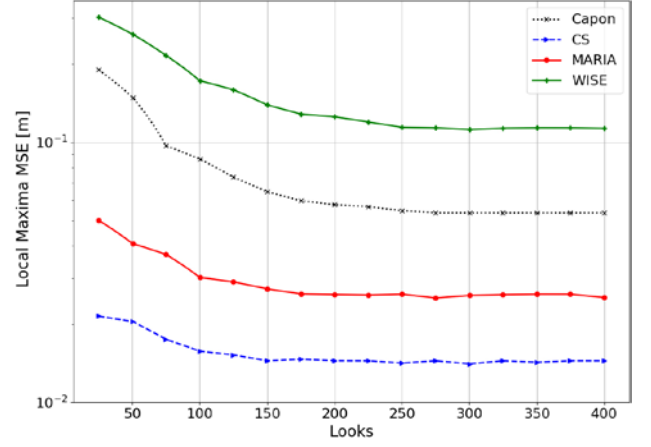


Fig. 3. MSE between the estimated and actual positions of the three most prominent local maxima for a spread value  $\sigma = 0.8$  m versus number of looks  $J$ . We refer to Gaussian distributions to model the backscattering. Each point is the average of 100 Monte-Carlo trials.

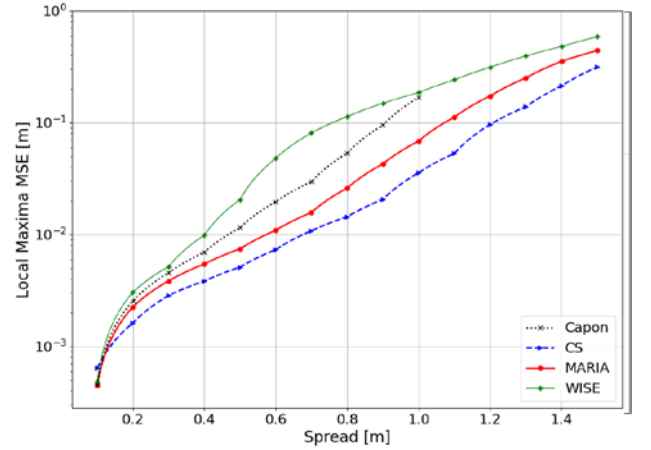


Fig. 4. MSE between the estimated and actual positions of the three most prominent local maxima versus spread value  $\sigma$ . We refer to Gaussian distributions to model the backscattering. Each point is the average of 100 Monte-Carlo trials.

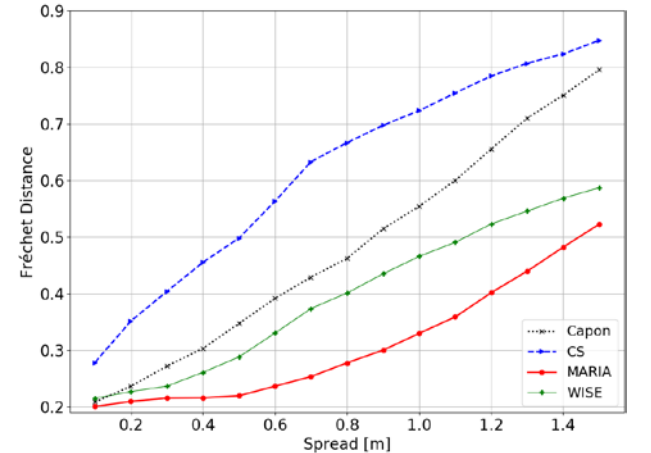


Fig. 5. Fréchet distance between the estimated and the theoretical backscattering profiles versus spread value  $\sigma$ . We refer to Gaussian distributions to model the backscattering. Each point is the average of 100 Monte-Carlo trials.

<sup>1</sup>Traditionally, the Fréchet distance is exemplified through the analogy of a man walking a dog. Suppose that a man is walking on one curve and a dog on another, both are allowed to adjust their speed but cannot move backwards. The Fréchet distance of the two curves is the minimum length of leash necessary to connect the man and the dog.

$$Fr(P, Q) = \inf_{\nu} \max_{x \in [0,1]} \|P(\nu(x)) - Q(\nu(x))\|, \quad (65)$$

where  $P, Q: [0,1] \rightarrow \mathbb{R}^2$  are parametrizations of the two curves and  $\nu: [0,1] \rightarrow [0,1]$  are continuous non-decreasing one-to-one functions, such that  $\nu(0)=0$ ,  $\nu(1)=1$  and  $\nu(0)=0$  and  $\nu(1)=1$ .

The values given in Fig. 5 were obtained using the algorithm presented in [33], which computes the Fréchet distance between two arbitrary curves by approximating them with polygonal curves.

As it can be observed through Fig. 4 and Fig. 5, the value of the MSE and Fréchet distance measurements increase as the spread value  $\sigma$  increases. This is expected, since the scatterers within each individual Gaussian distribution start mixing together, which makes the PSP more challenging to estimate.

The Capon's curve in Fig. 4 stops at the spread value of  $\sigma = 1.0$  m, since later on such solver is not capable to distinguish the presence of three local maxima.

Fig. 4 shows that the competing state-of-the-art CS [10], [11] and MARIA super-resolved techniques perform better on the retrieval of the location of the corresponding local maxima. For the reported numerical examples, CS retrieves the more accurate estimates of the local maxima, followed by the MARIA technique. Since the simulated scatterers follow well-defined Gaussian distributions, it is expected that MARIA performs better in this subject than the introduced WISE approach.

Although, CS recovers more accurate positions of the local maxima, it incurs much more computation time in contrast to the other competing methods, as depicted in Table IV. This is due to the iterative nature of the method and due to the non-availability of adapted efficient convex optimization algorithms [11, Section V]. Specifically, the CVXPY software library [34] has been employed to solve the involved convex optimization problem in (64).

Additionally, Fig. 5 shows that MARIA and WISE provide more information on the actual outline of the simulated backscattering vertical profiles, in contrast to the other CS competing technique.

Being WISE an extension of the Capon method [3], [7] – [9], we now validate the feature-enhancement characteristics of this novel technique via two individual representative cases: (i) A volumetric target with spread of  $\sigma = 0.8$  m (as depicted in Fig. 2a), for which Capon beamforming still distinguishes the presence of three local maxima; (ii) A volumetric target with spread of  $\sigma = 1.3$  m (as shown in Fig. 2b), for which Capon beamforming is not able to discriminate the presence of three local maxima. The comparison against the TomoSAR-adapted competing methods, MSF beamforming in (10), Capon in (62), CS in (64) and MARIA in (60), is presented through Fig. 6 and Fig. 7, for the respective cases. Notice that the true positions of the phase-centers, for each individual distribution, are pointed out through three black circles.

TABLE IV  
AVERAGE PROCESSING TIME<sup>1</sup>

TomoSAR Focusing Technique	Processing Time in Seconds
Capon	0.365164
CS	27.010265
MARIA	3.420420
WISE	3.296872

<sup>1</sup>The average processing time refers to the retrieval of a single profile, performed in an Intel® Xeon® X7560 CPU at 2.27 GHz, using a single thread.

Since the addressed MARIA and WISE statistical regularization approaches involve several matrix-vector operations, which include matrix inversions, we do not expect to preserve radiometric accuracy as the standard Fourier beamforming [1] does. Thus, we refer to a (0 to 1) normalized pseudo-power in the plots given afterward.

Fig. 4 shows that, in contrast to WISE, Capon recovers more accurate positions of the previously specified local maxima for spread values less than  $\sigma = 1.0$  m. However, WISE retrieves better Fréchet distances, as seen in Fig. 5.

As it can be observed via Fig. 6(b) and Fig. 7(b), the concave up curves that separate the Gaussian distributions of the actual PSP (depicted in Fig. 2) are more difficult to follow for Capon, especially when  $\sigma$  increases. Consequently, the computed Fréchet distances for Capon worsen as the spread value  $\sigma$  increases.

At this point, we can identify that the MARIA and WISE techniques provide particular advantages when coping with volumetric targets, modelled as an assembly of random scatterers with a certain spread (see Fig. 2). Specifically, both, WCF-based and ML-inspired approaches perform resolution-enhancement to a first estimate of the continuous PSP (retrieved using Capon). This is especially noticeable when the spread value  $\sigma$  is such that, the conventional Capon beamforming approach is not able to differentiate the existence of multiple (three) local maxima. Also, we observe that both methods recover an outline of the backscattering vertical profile more similar to the expected (theoretical) one, in contrast to the other competing techniques, which is corroborated through the Fréchet distance measurements of Fig. 5.

Next, we make use of a different distribution to model the echoes coming from the scatterers, gathered on each simulated data covariance matrix  $\mathbf{Y}$ . The scatterers have equal reflectivity and are displaced along the height direction. Particularly, we refer to three Rayleigh distributions located at  $z_1 = -3$  m,  $z_2 = 0$  m and  $z_3 = 4$  m, correspondingly. A varying spread (scale parameter)  $0.8 \text{ m} \leq \sigma \leq 2.3 \text{ m}$  (Fig. 8 shows two representative cases) and a noise power spectral density  $N_0 = 0.01$  are considered.

The same TomoSAR focusing techniques as before are applied to recover the PSP for the respective cases. A high number of looks ( $J = 350$ ) is chosen, for which Capon beamforming presents enhanced performance, as verified previously.

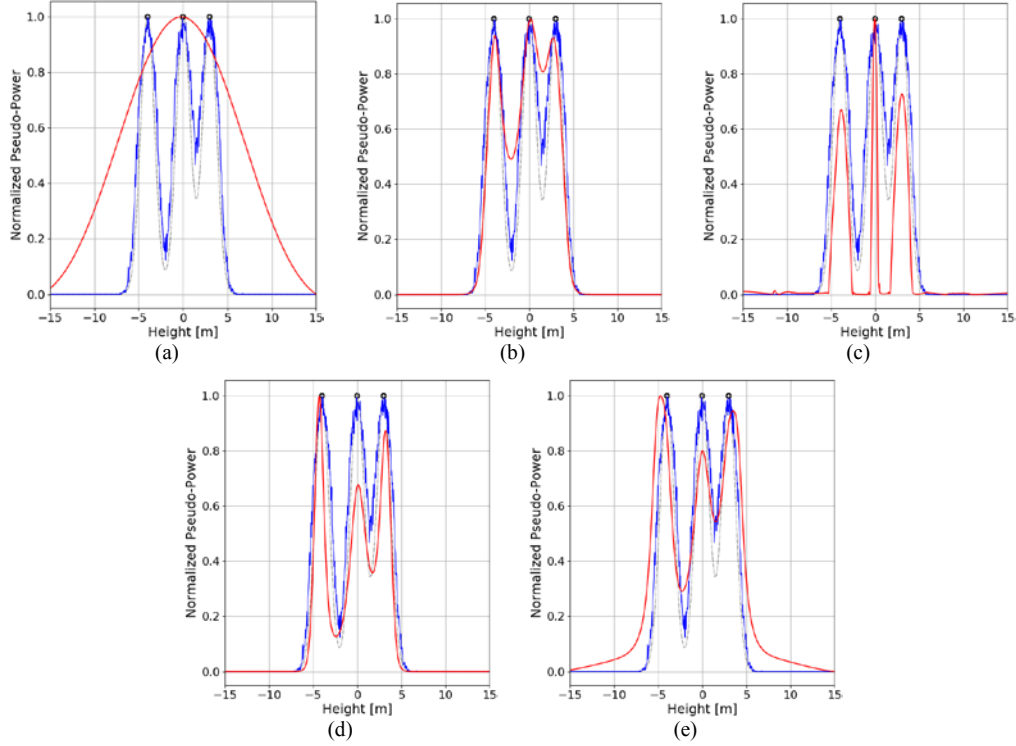


Fig. 6. **Blue:** Average among looks of the simulated signal. We refer to three Gaussian distributions, each one with spread  $\sigma = 0.8$  m. The true positions of the respective phase-centers, for each individual distribution, are pointed out through three black circles. **Red:** Vertical profile recovered after applying (a) MSF beamforming [3], [19]; (b) Capon [3], [7] – [9]; (c) wavelet-based CS [10], [11] with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition ( $\tau_1 = \tau_2 = 0.5$ ); (d) the MARIA technique in (44) after  $I = 6$  iterations ( $N_0 = 0.01$ ); (e) the WISE method in (54) after  $I = 5$  iterations ( $N_0 = 0.01$ ).

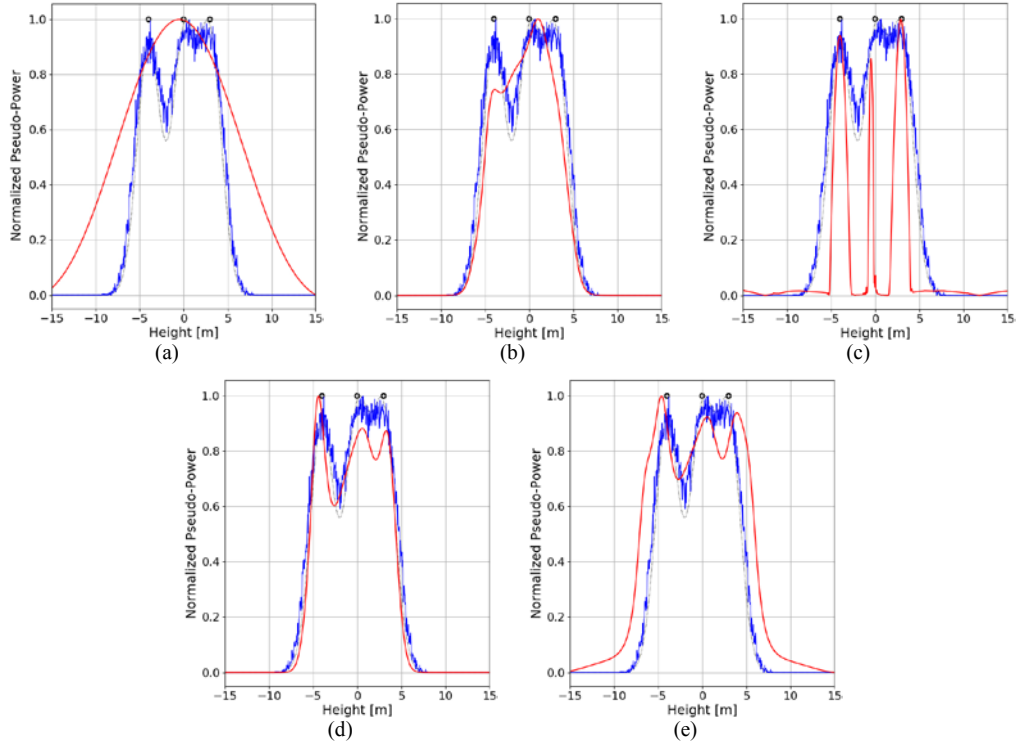


Fig. 7. **Blue:** Average among looks of the simulated signal. We refer to three Gaussian distributions, each one with spread  $\sigma = 1.3$  m. The true positions of the respective phase-centers, for each individual distribution, are pointed out through three black circles. **Red:** Vertical profile recovered after applying (a) MSF beamforming [3], [19]; (b) Capon [3], [7] – [9]; (c) wavelet-based CS [10], [11] with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition ( $\tau_1 = \tau_2 = 0.5$ ); (d) the MARIA technique in (44) after  $I = 10$  iterations ( $N_0 = 0.01$ ); (e) the WISE method in (54) after  $I = 13$  iterations ( $N_0 = 0.01$ ).



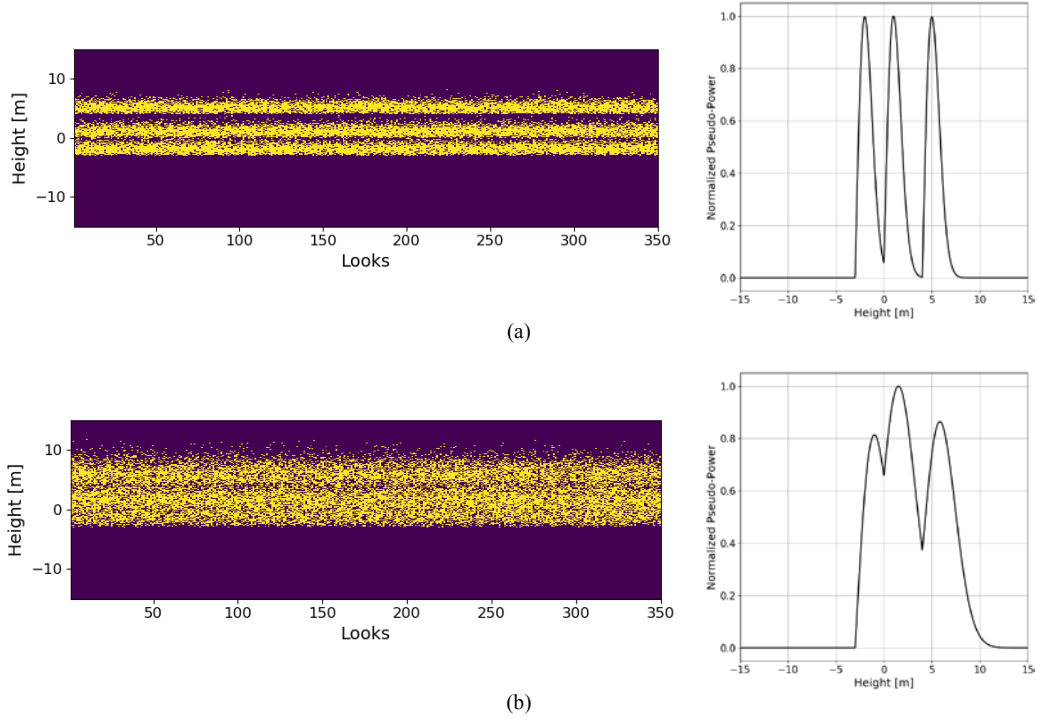


Fig. 8. Distribution of 300 scatterers displaced along the height direction, following three Rayleigh distributions among  $J = 350$  independent looks, located at  $z_1 = -3$  m,  $z_2 = 0$  m and  $z_3 = 4$  m correspondingly, and with spread (scale parameter) of (a)  $\sigma = 1.0$  m and (b)  $\sigma = 2.0$  m, respectively. The theoretical distribution, for each case, is presented at the right side.

Afterward, we compare the retrieved three most prominent local maxima against the expected (theoretical) local maxima. Keep in mind that the mean value (phase-center) of a Rayleigh distribution is not located at its maximum.

For such a case, Fig. 9 compares the MSE between the estimated and actual positions of the three most prominent peaks, retrieved using the competing methods (Capon, CS and MARIA) and the proposed WISE for different spread values  $\sigma$ . Fig. 10 compares the Fréchet distance [32], [33], between the estimated and actual (theoretical) vertical profiles, correspondingly.

Observe via Fig. 9 and Fig. 10 that, as previously, the MSE and Fréchet distance measurements increase as the scatterers,

within each individual Rayleigh distribution, start mixing together. Notice also that all curves in Fig. 9 and Fig. 10 grow slower as the spread value  $\sigma$  increases, in contrast to Fig. 4 and Fig. 5, respectively.

The Capon's curve in Fig. 9 stops at the spread value of  $\sigma = 1.4$  m, since later on such technique is not able to discriminate the presence of three local maxima.

As expected, MARIA recovers worse MSE and Fréchet distance measurements in comparison to the results gotten from the scatterers following Gaussian distributions. On the other hand, CS remains as the technique with better performance on the retrieval of the location of the corresponding local maxima.

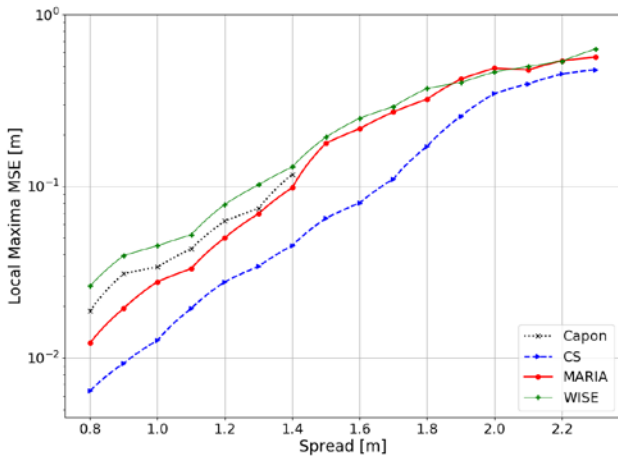


Fig. 9. MSE between the estimated and actual positions of the three most prominent local maxima versus spread value  $\sigma$ . We refer to Rayleigh distributions to model the backscattering. Each point is the average of 100 Monte-Carlo trials.

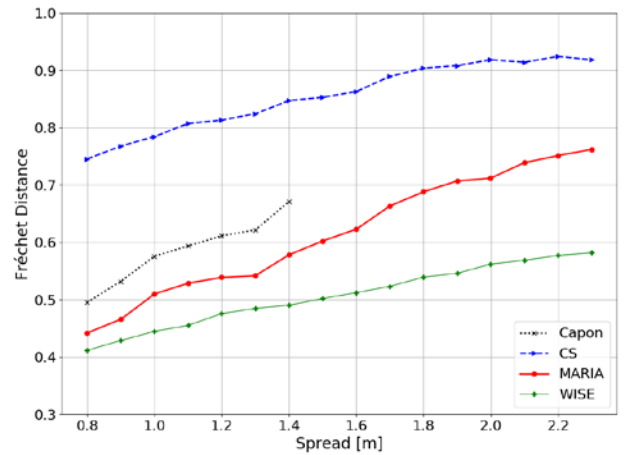


Fig. 10. Fréchet distance between the estimated and the theoretical backscattering profiles versus spread value  $\sigma$ . We refer to Rayleigh distributions to model the backscattering. Each point is the average of 100 Monte-Carlo trials.

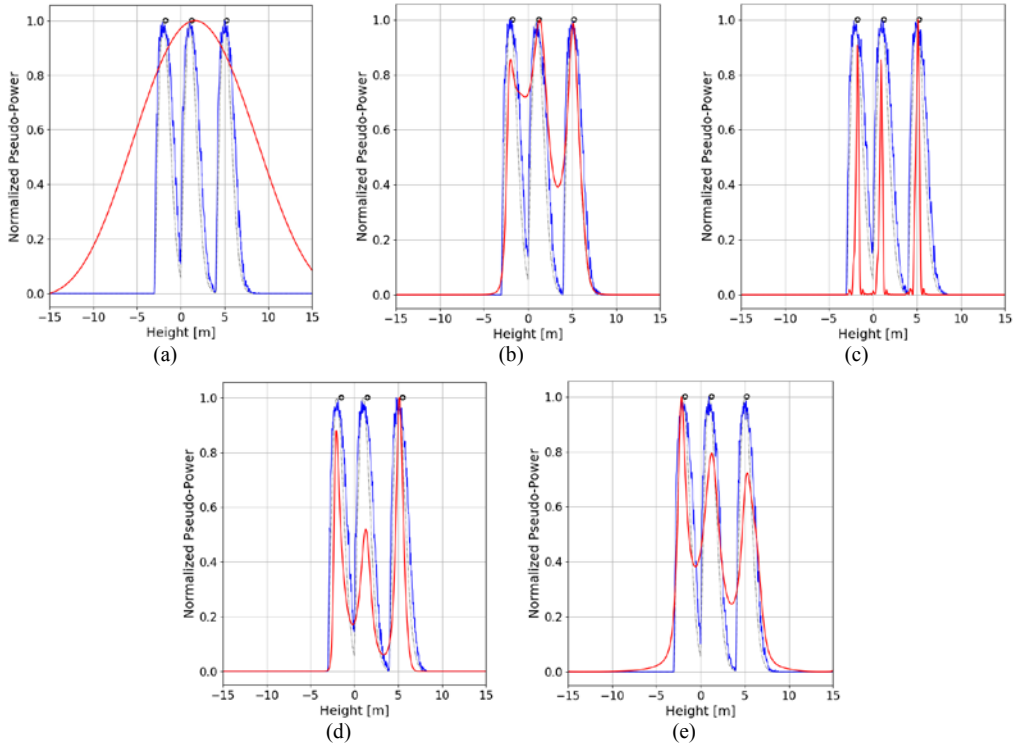


Fig. 11. **Blue:** Average among looks of the simulated signal. We refer to three Rayleigh distributions, each one with spread  $\sigma = 1.0$  m. The true positions of the respective phase-centers, for each individual distribution, are pointed out through three black circles, namely: -1.75 m, 1.25 m and 5.25 m, respectively. **Red:** Vertical profile recovered after applying (a) MSF beamforming [3], [19]; (b) Capon [3], [7] – [9]; (c) wavelet-based CS [10], [11] with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition ( $\tau_1 = \tau_2 = 0.5$ ); (d) the MARIA technique in (44) after  $I = 5$  iterations ( $N_0 = 0.01$ ); (e) the WISE method in (54) after  $I = 6$  iterations ( $N_0 = 0.01$ ).

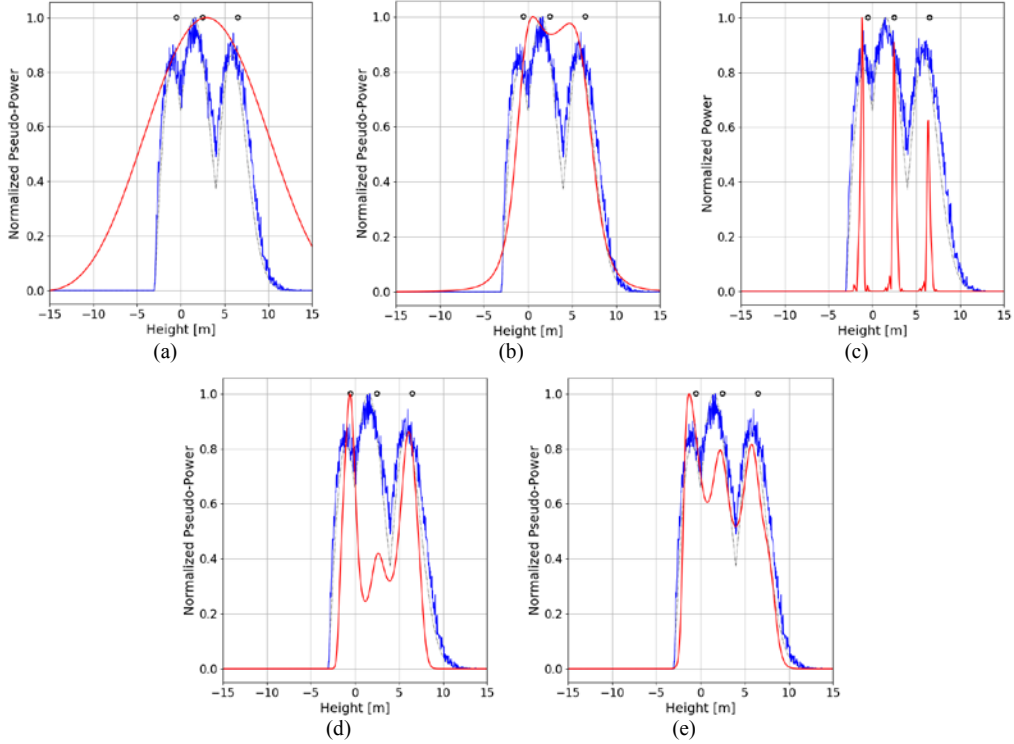


Fig. 12. **Blue:** Average among looks of the simulated signal. We refer to three Rayleigh distributions, each one with spread  $\sigma = 1.0$  m. The true positions of the respective phase-centers, for each individual distribution, are pointed out through three black circles, namely: -0.50 m, 2.50 m and 6.5 m, respectively. **Red:** Vertical profile recovered after applying (a) MSF beamforming [3], [19]; (b) Capon [3], [7] – [9]; (c) wavelet-based CS [10], [11] with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition ( $\tau_1 = \tau_2 = 0.5$ ); (d) the MARIA technique in (44) after  $I = 10$  iterations ( $N_0 = 0.01$ ); (e) the WISE method in (54) after  $I = 9$  iterations ( $N_0 = 0.01$ ).



The super-resolved WISE technique presents a similar behavior as earlier, showing more independence on how the scatterers are distributed along the height direction. This time, MARIA and WISE retrieve more similar results on the estimation of the local maxima positions, which we attribute to a poorer performance of MARIA rather than a significant improvement of WISE. Also, WISE presents the best Fréchet distance measurements, which translates into recovering an outline of the backscattering vertical profile more similar to the expected (theoretical) one.

Finally, Fig.11 and Fig. 12 present, as before, the comparison of WISE in (61) against the most prominent non-parametric competing methods (MSF in (10), Capon in (62), CS in (64) and MARIA in (60)) for two individual representative cases: (i) A volumetric target with spread of  $\sigma=1.0$  m (as depicted in Fig. 8a), for which Capon beamforming still distinguishes the presence of three local maxima; (ii) A volumetric target with spread of  $\sigma=2.0$  m (as shown in Fig. 8b), for which Capon beamforming is not able to discriminate the presence of three local maxima. Notice that the true positions of the phase-centers, for each individual distribution, are pointed out through three black circles.

The presented analyses confirm what is expected from the behavior of WISE, in comparison, especially, to the closest most competing MARIA method: the use of WISE is more suitable for scenarios where the observed data comes from scatterers following an explicitly non Gaussian distribution.

In addition, we can come to an important conclusion, explained next. Both regularization techniques, MARIA and WISE, require of the proper setting of  $N_0$ , which according to our model in (5), represents the power spectral density of the white noise power [3]. Nonetheless, in practice, this factor assumes the role of a regularization parameter.

For the given simulations we set  $N_0 = 0.01$ , as originally defined to model the backscattering. However, looking at Fig. 4 and Fig. 9, the fact that WISE retrieves worse results than its input Capon, means that the WISE solutions are not optimally regularized. This indicates that the value of  $N_0$  is not properly chosen. Furthermore, being WISE an iterative procedure, it is likely that this factor needs to be updated at each iteration.

We presume then that WISE can improve its performance by fine-tuning  $N_0$  within formula (61). Having a similar structure, this also applies to MARIA in (60).

The optimal selection of the regularization parameter  $N_0$  is out of the scope of this article, since it requires of dedicated work. Such problem is to be tackled in a general way, independently of the regularization approach to be used, as suggested in [35].

Meanwhile, for the experiments reported next, we treat  $N_0$  now as a user specified regularization parameter, defined for each particular method. Notice that the assigned values are indicated in the figure captures.

## VIII. EXPERIMENTAL RESULTS

The reported experimental results are obtained via processing L-band (0.226 m wavelength) data of the German Aerospace Center (DLR), acquired by the F-SAR airborne

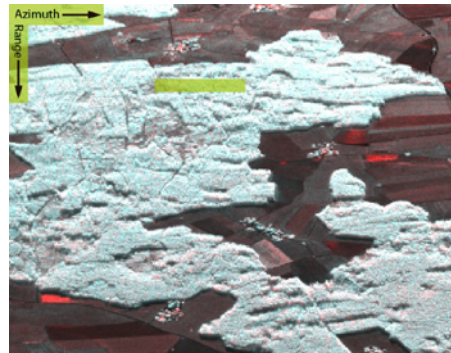


Fig. 13. Polarimetric SAR image of the forested test site in Froschham, Germany, 2017 (Red:  $HH$ ; Blue:  $HV$ ; Green:  $VH$ ). The area of interest is located within the yellow rectangle along azimuth.

sensor at a nominal altitude above ground of 3720 m, over the forested test site of Froschham, Germany, in 2017. The acquisition constellation spans a horizontal synthetic aperture of 70 m comprising 15 tracks approximately evenly distributed. Thus, we make use of a stack of 15 single-look complex SAR images (Fig. 13 shows one image out of the stack) properly co-registered and phase flattened.

In practice, regions with high air traffic, such as the one considered in this study, demand to operate over only one height level, so that the flight space is not blocked. Therefore, the acquisition constellation rather spans a horizontal synthetic aperture.

The tomographic reconstruction is conducted as follows: several azimuth positions at a fixed range distance are selected as indicated by the yellow rectangle in Fig. 13.

For the chosen area, with a slant-range distance of about 4000 m between the master track and the targets, and a look angle of about  $40^\circ$ , the Fourier resolution [4] is 5.4 m when we consider all the passes.

To demonstrate the feature-enhancement capabilities of the WCF-inspired regularization approach, in comparison to the other treated state-of-the-art competing techniques, we employ only half ( $L = 7$ ) of the existing tracks, meaning a Fourier resolution of 12.6 m. Then, the retrieved results are validated with the tomogram obtained after applying MSF beamforming using all the available passes ( $L = 15$ ), as shown in Fig. 14. Polarization  $HV$  is chosen in order to have fewer contributions from the ground surface and to emphasize the whole vegetation structure.

Besides of showing the tomograms for the different cases, Fig. 14 also presents a comparison between the retrieved vertical profiles for a chosen azimuth position. We select the azimuth position at 155 m, since the presence of multiple phase-centers is detected in this zone. For each tomogram, the corresponding vertical profile is shown at right hand. The location among the tomogram, from where the vertical profile is taken, is depicted through a green vertical line crossing the tomogram at the chosen azimuth position (155 m). For an easy assessment, the outline of the vertical profile retrieved using MSF with all passes, is super-imposed onto the other corresponding plots, depicted via a dashed red line.

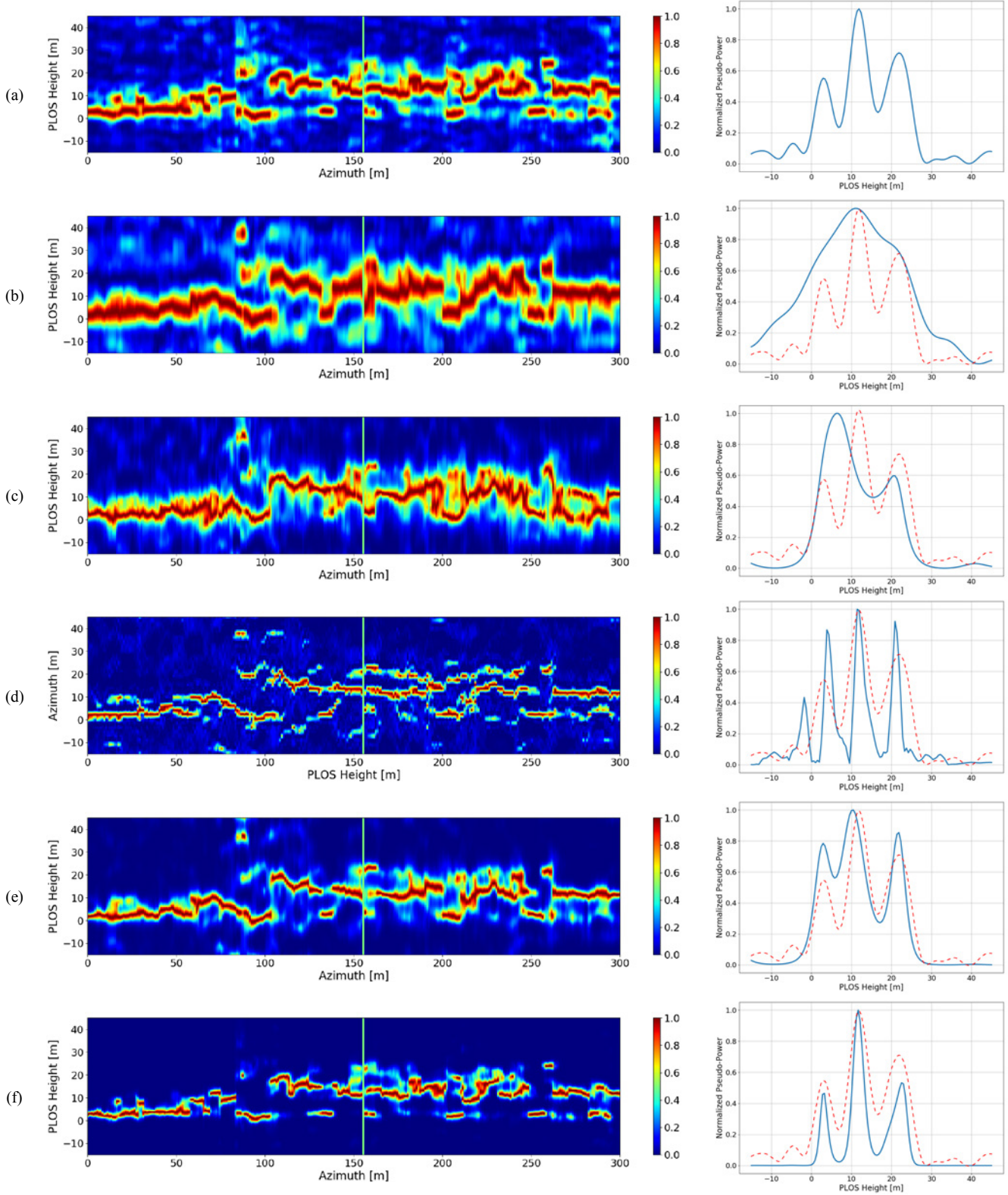


Fig. 14. Retrieved  $HV$  tomograms and the corresponding vertical profile at the azimuth position 155 m, after applying: (a) MSF beamforming in (10) with  $L=15$  tracks; (b) MSF beamforming with  $L=7$  tracks; (c) Capon in (62) with  $L=7$  tracks; (d) wavelet-based CS in (64) with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition and  $L=7$  tracks ( $\tau_1 = \tau_2 = 0.5$ ); (e) the MARIA technique in (60) after  $I = 8$  iterations with  $L=7$  tracks ( $N_0 = 0.0045$ ); (f) the WISE method in (61) after  $I = 9$  iterations with  $L=7$  tracks ( $N_0 = 0.075$ ). We use a Boxcar filter with a  $7 \times 7$  pixels window to perform multi-looking on the array of data covariance matrices. The retrieved power is presented in a normalized (0 to 1) linear scale. For an easy assessment, the outline of the vertical profile retrieved using MSF with all passes, is superimposed onto the other corresponding plots, depicted via a dashed red line.

The attainable resolution of the conventional TomoSAR-adapted DOA-based non-parametric methods, such as the MSF and Capon techniques, highly depends on the tomographic aperture [3], [4], [6]. Due to this constraint, as observed in Fig. 14(b) and 14(c), these techniques are not able to resolve two arbitrarily close scattering contributions with a reduced (limited) number of tracks.

More sophisticated approaches as CS [10], [11] and MARIA [13], or the novel introduced WISE method, are capable to discriminate the contributions from the several phase-centers, since such methods are not constrained to the resolution of the sensor geometry, as they do not involve a conventional inversion of the spectrum [12]. This is validated through Fig. 14(d), 14(e) and 14(f), correspondingly. Nonetheless, as it is shown in Fig. 14(d), the wavelet-based CS technique retrieves higher ambiguity levels, which may translate into false detections.

Fig. 14(e) and 14(f) depict how the ML-inspired and the novel WCF-based regularization approaches, perform resolution-enhancement, with suppression of artifacts and ambiguity levels reduction, to an initial estimate of the continuous PSP, retrieved using the Capon beamforming technique, given by Fig. 14(c). Such statistical regularization approaches allow for the separation of multiple scattering contributions within each resolution cell. After convergence, they are able to recover an accurate location of the corresponding closely spaced local maxima, which is validated with the tomogram retrieved using MSF beamforming with all the available passes (Fig. 14(a)). Being non-parametric, it is important to remark, that both regularization approaches do not require any a priori information about the number of backscattering sources that constitute the illuminated area, in contrary to other super-resolved parametric estimation techniques as MUSIC [3], [4] or ESPRIT [8].

In order to show the stability of the proposed method, we present the superposition of the vertical profiles along two (non-homogeneous) areas within the tomograms in Fig. 14. For each tomogram described in Fig. 14, Fig. 15 shows the super-imposed vertical profiles comprising the azimuth positions between 125 m and 175 m, whereas Fig. 16 considers the azimuth positions between 225 m and 275 m.

An important factor to take into account is the mixture of sources consequence of spatial multi-looking. In contrast to the numerical examples presented before, where a high number of independent realizations (looks) are simulated, in practice, TomoSAR is customarily treated as an ergodic process, meaning that its statistical properties are deduced from a single random realization. The latter usually translates into achieving multi-looking through the averaging of adjacent values among the set of data covariance matrices  $\mathbf{Y}$ , i.e., using a Boxcar filter.

The introduced WISE technique, along with MARIA, requires of a high number of looks for a better performance, since, by constructions, the first estimate  $\hat{\mathbf{b}}^{[0]}$  of the PSP  $\mathbf{b}$  is retrieved using Capon beamforming. The latter may lead to the spatial mixture of sources.

In the experimental results given in Fig. 14, we perform multi-looking using a Boxcar filter with a  $7 \times 7$  pixels

window on the array of data covariance matrices. Afterward, in order to show the mixture of sources as an effect of spatial multi-looking, we use now a Boxcar filter with a  $15 \times 15$  pixels window to recover the tomograms in Fig. 17.

Fig. 17 shows the same tomograms retrieved previously in Fig. 14, but using a different window size ( $15 \times 15$  pixels) to perform multi-looking via Boxcar filtering. The tomograms shown in Fig. 17 are recovered from the exactly same region as in Fig. 14, depicted by the yellow rectangle in Fig. 13.

As one can observe, the tomographic slices in Fig. 17 present a reduced resolution, in contrast to the tomograms given in Fig. 14, because of the spatial averaging. However, the presence of multiple local maxima is still detected, as shown through the comparison between the retrieved vertical profiles at the chosen azimuth position (155 m), for Fig. 14 and Fig. 17, respectively.

The simulation results in the previous section show that MARIA and the novel proposed WISE provide more information on the outline of the backscattering vertical profiles. For real-world scenarios, the WISE method in (61) is expected to perform better in this subject, since such approach is more flexible to the pdf's Gaussian assumption of the observed data (which may be characterized with a different pdf), in contrast to the MARIA (60) technique.

The latter is especially noticeable through the reported experimental results in Fig. 17(e) and 17(f), where we can observe that the retrieved vertical profiles of the MARIA approach are composed of Gaussian like reconstructions, and the retrievals of the WISE technique tend to be more similar to the results gotten using MSF beamforming [3], [17], with all the available tracks, as shown in Fig. 17(a). Our impression is then that high multi-looking also aids WISE to this subject.

Consequently, in order to achieve high multi-looking, preventing the spatial mixture of sources, rather than using a Boxcar filter, we recommend the use of non-local spatially adaptive filtering methods, which enhance the estimation of the covariance matrices, improving the scatterer separation in layover areas thanks to their smoothing and edge-preserving properties. However, this topic is out of the scope of this article, but the reader may refer to [36], [37].

The work presented in [38] explains how to characterize the forest structure from the vegetation layers that compose it, reflected as local maxima in the tomographic profiles. The presence or absence of multiple maxima in certain zones, impacts directly on the forest structure measurements. Thus, enhanced resolution is desired to detect the corresponding closely spaced local maxima. As it has been corroborated, MARIA and WISE ease such efforts by providing enhanced-resolution and by retrieving an accurate location of the peaks.

Recall that, as other super-resolved techniques, the addressed MARIA and WISE statistical regularization approaches trade radiometric accuracy for enhanced-resolution. This is clearly noticeable, since both iterative procedures involve several matrix-vector operations, which include matrix inversions. Consequently, for comparison purposes, we refer to a (0 to 1) normalized pseudo-power in the plots given in Fig. 14 - Fig. 17.



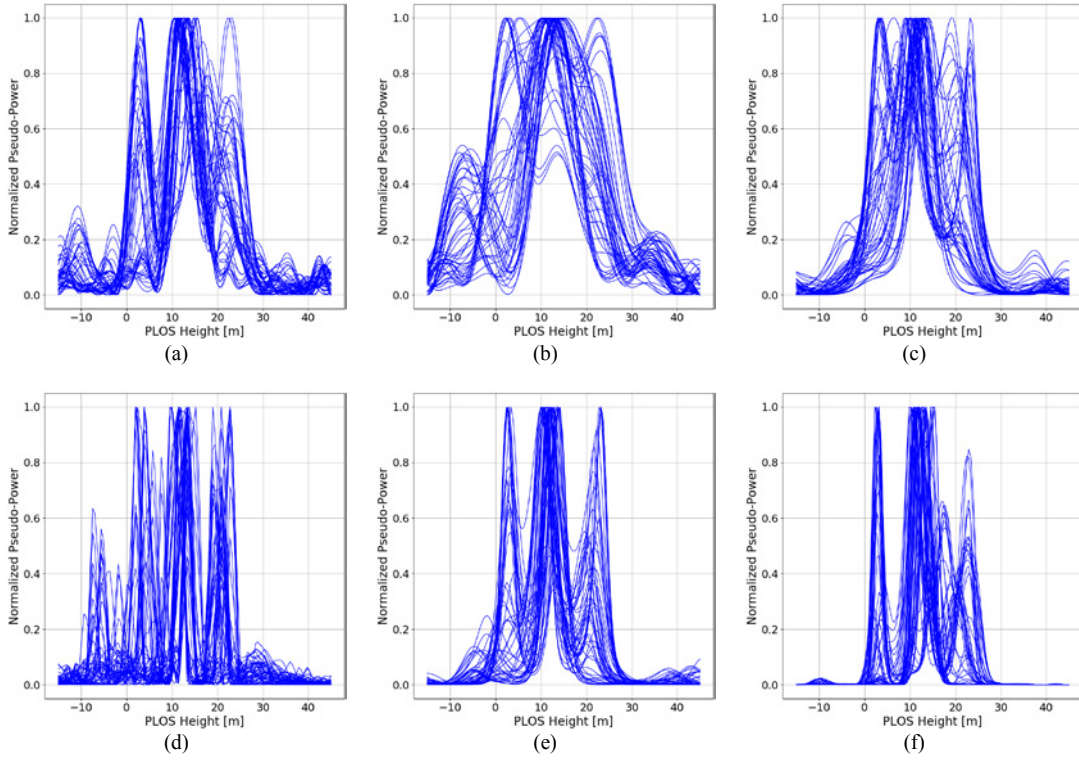


Fig. 15. Superposition of fifty vertical profiles comprising the azimuth positions between 125 m and 175 m within the tomograms in Fig. 14: (a) MSF in (10) with  $L=15$  tracks; (b) MSF with  $L=7$  tracks; (c) Capon in (62) with  $L=7$  tracks; (d) wavelet-based CS in (64) with  $L=7$  tracks; (e) MARIA in (60) with  $L=7$  tracks; WISE in (61) with  $L=7$  tracks.

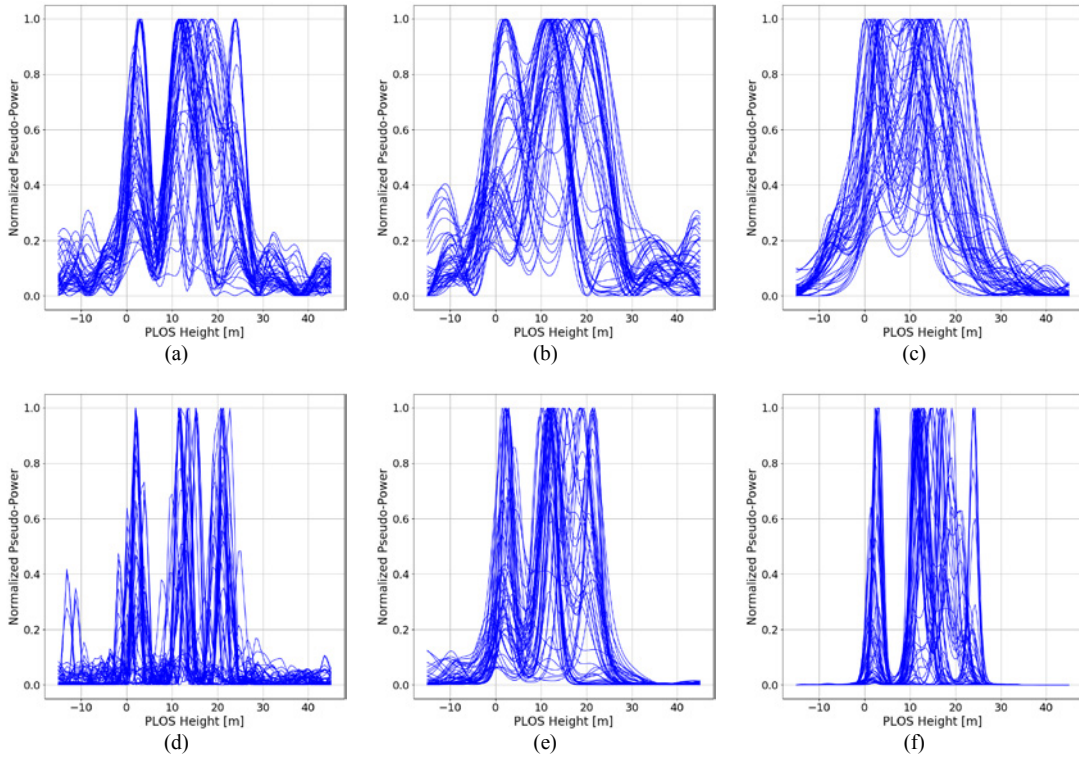


Fig. 16. Superposition of fifty vertical profiles comprising the azimuth positions between 225 m and 275 m within the tomograms in Fig. 14: (a) MSF in (10) with  $L=15$  tracks; (b) MSF with  $L=7$  tracks; (c) Capon in (62) with  $L=7$  tracks; (d) wavelet-based CS in (64) with  $L=7$  tracks; (e) MARIA in (60) with  $L=7$  tracks; WISE in (61) with  $L=7$  tracks.

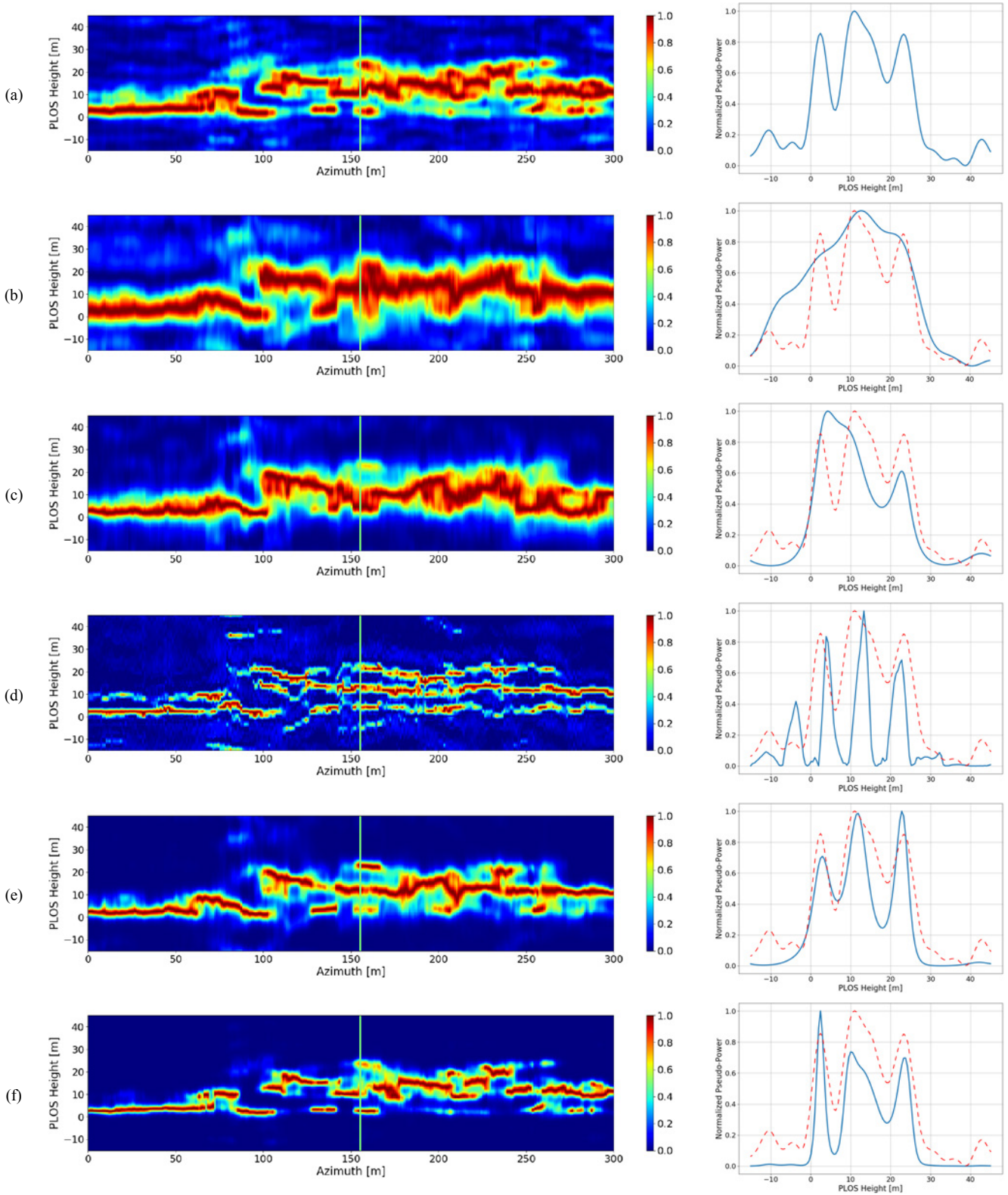


Fig. 17. Retrieved  $HV$  tomograms and the corresponding vertical profile at the azimuth position 155 m, after applying: (a) MSF beamforming in (10) with  $L=15$  tracks; (b) MSF beamforming with  $L=7$  tracks; (c) Capon in (62) with  $L=7$  tracks; (d) wavelet-based CS in (64) with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition and  $L=7$  tracks ( $\tau_1 = \tau_2 = 0.5$ ); (e) the MARIA technique in (60) after  $I = 8$  iterations with  $L=7$  tracks ( $N_0 = 0.0045$ ); (f) the WISE method in (61) after  $I = 9$  iterations with  $L=7$  tracks ( $N_0 = 0.075$ ). These tomograms correspond to the exactly same region as in Fig. 14, depicted by the yellow rectangle in Fig. 13, but using a  $15 \times 15$  pixels window to perform multi-looking via Boxcar filtering. The retrieved power is presented in a normalized (0 to 1) linear scale. For an easy assessment, the outline of the vertical profile retrieved using MSF with all passes, is superimposed onto the other corresponding plots, depicted via a dashed red line.

For other methodologies and applications, normalizing each vertical profile with respect to one could be misleading, giving equal importance to mechanisms with different reflectivity. Subsequently, Fig. 18 shows the same tomograms as in Fig. 14, but normalizing them with respect to the trace (span) of the covariance matrices and in a dB scale. We leave to the reader the suitability of the presented regularization techniques to the particular applications.

With the addressed considerations, the introduced method is presented as a feasible candidate for a practical implementation in the perspective of future airborne and spaceborne TomoSAR missions.

## IX. CONCLUDING REMARKS

With the aim of solving the ill-conditioned nonlinear TomoSAR inverse problem, this paper introduces WISE, a novel non-parametric statistical regularization approach based on the WCF criterion. The novel technique relaxes the pdf's Gaussian assumption of the observed data, in contrast to the ML-inspired approaches, as MARIA, which consider Gaussian statistics.

The WISE algorithm allows for its application to man-made (urban) and natural scenes; nevertheless, since we refer to a model better adapted for the representation of natural environments, we expect improved performance in such scenarios, i.e., rough surfaces and volumetric targets as forests.

First, we present the mathematical background of the most prominent regularization techniques, from which WISE is derived. These techniques have been assembled into two main groups: (i) deterministic descriptive regularization techniques (CLS and WCLS), inspired by the Tikhonov's regularization theory; and (ii) statistical regularization methods (BMR, ML and WCF), in the context of the statistical decision-making theory. As a consequence of this analysis, WISE is yielded as an extension of Capon beamforming and with a similar structure to MARIA, which is derived here in a more convenient fashion.

At this point, we identify that WISE is a solution dependent technique, which requires of a priori information on the PSP. Furthermore, it involves several matrix-vector operations, including matrix inversions; therefore, keep in mind that the radiometric accuracy is not preserved.

Next, the iterative adaptive implementation of WISE is given. This implementation is considered as a sub-optimal version of the WCF solver, since the actual PSP is unknown and needs to be firstly estimated. For such purpose, by constructions, the conventional Capon beamforming technique is employed.

Capon beamforming is widely used to perform focusing of the TomoSAR data, since it provides enhanced resolution when a high number of looks is processed. Recall that, due to the involved inversions of the data covariance matrices, the rank-deficiencies of such matrices must be avoided.

Be aware that, technically, multi-looking is commonly achieved through the averaging of adjacent values via Boxcar filtering, which may lead to the spatial mixture of sources. For

this reason, the use of non-local spatially adaptive filtering methods is recommended instead.

Afterward, the capabilities of WISE are demonstrated through numerical examples, in comparison to the most prominent non-parametric competing techniques MSF, Capon, CS and MARIA.

From such analyses, we observe that MSF and Capon are not always able to discriminate the presence of multiple local maxima due to their inferior attainable resolution.

Also, when the PSP is properly modelled using Gaussian statistics, MARIA, the closest competing method, retrieves better results than WISE on the retrieval of the Fréchet distance measurements and local maxima positions. However, the performance of MARIA worsens when we consider different (non-Gaussian) statistics.

Recall that we use the Fréchet distance to compare the recovered and expected (theoretical) outline of the backscattering vertical profiles.

In general, CS has a better performance on the retrieval of the local maxima positions, but with a higher computational burden. On the other hand, CS presents the worst Fréchet distance measurements.

As expected, WISE shows more independence on how the scatterers are distributed along the height direction. It retrieves similar results to MARIA on the estimation of the local maxima positions, when non-Gaussian statistics are considered. Moreover, for such a case, WISE retrieves the best Fréchet distance measurements.

Next, from the inspection of the simulation results, we infer that factor  $N_0$ , which represents the power spectral density of the white noise power in the TomoSAR signal model, should be rather treated as a user specified regularization parameter. We presume that WISE can improve its performance by fine-tuning  $N_0$ .

One can conclude then, that the use of WISE is more suitable for scenarios where the pdf of the observed data is explicitly non Gaussian. Such could be the case of non-uniform sparse forests as tropical forests, known for its dense canopies of vegetation that form multiple layers; or the subsurface imaging of ice structures as glaciers, which normally have different ice thickness among the covered area, along with different subsurface and subglacial conditions, and meltwater discharge rates.

Later on, the behavior of WISE and the other competing techniques is validated using real data. For this purpose, we refer to a natural environment, namely, a forested scene.

As it has been verified, WISE presents the following advantages: (i) being non-parametric, it is not necessary to know beforehand the number of sources sharing the same resolution cell; (ii) it recovers resolution-enhanced tomographic profiles with a reduced (limited) number of passes; (iii) it retrieves reduced ambiguity levels; (iv) it performs suppression of artifacts; (v) it provides meaningful information on the outline of the actual backscattering vertical profiles; (vi) it requires much less computation time in contrast to CS.



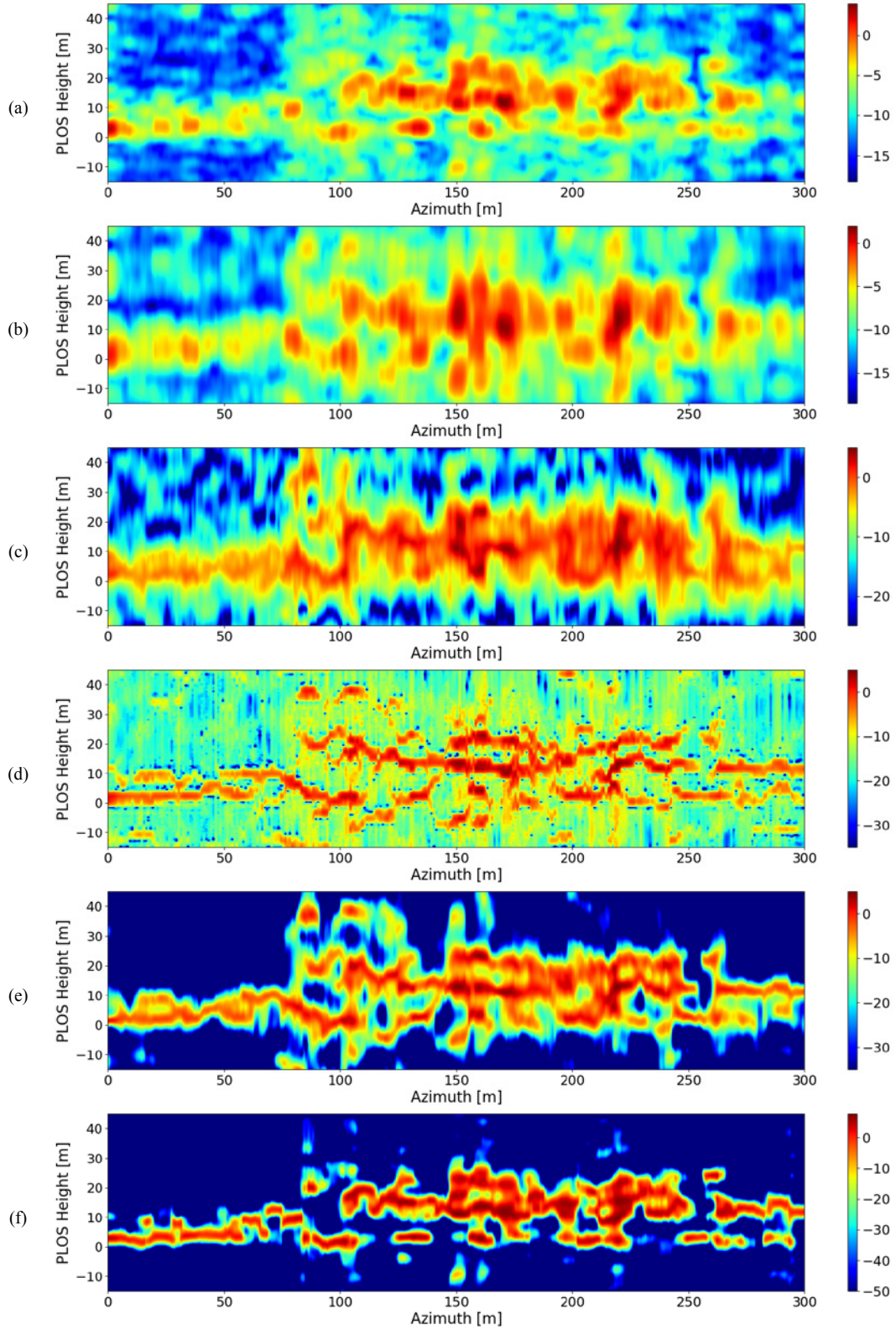


Fig. 18. Retrieved  $HV$  tomograms after applying: (a) MSF beamforming in (10) with  $L=15$  tracks; (b) MSF beamforming with  $L=7$  tracks; (c) Capon in (62) with  $L=7$  tracks; (d) wavelet-based CS in (64) with a sparsifying basis based on the Symlets 4 wavelet family with 3 levels of decomposition and  $L=7$  tracks ( $\tau_1 = \tau_2 = 0.5$ ); (e) the MARIA technique in (60) after  $I = 8$  iterations with  $L=7$  tracks ( $N_0 = 0.0045$ ); (f) the WISE method in (61) after  $I = 9$  iterations with  $L=7$  tracks ( $N_0 = 0.075$ ). We use a Boxcar filter with a  $7 \times 7$  pixels window to perform multi-looking on the array of data covariance matrices. These tomograms correspond to the exactly same region as in Fig. 14, depicted by the yellow rectangle in Fig. 13. The retrieved power is presented in a dB scale; it has been normalized with respect to the trace (span) of the corresponding covariance matrices.

## APPENDIX

This section gathers all the differentiation rules employed to derive the addressed regularization approaches. Particularly, the CLS (13) and WCLS (20) solvers utilize the following formulas [24, Appendix A.7],

$$\frac{\partial}{\partial \mathbf{u}} \mathbf{v}^+ \mathbf{u} = \mathbf{0}, \quad (\text{A.1})$$

$$\frac{\partial}{\partial \mathbf{u}} \mathbf{u}^+ \mathbf{v} = 2\mathbf{v}, \quad (\text{A.2})$$

$$\frac{\partial}{\partial \mathbf{u}} \mathbf{u}^+ \mathbf{V} \mathbf{u} = 2\mathbf{V} \mathbf{u}. \quad (\text{A.3})$$

Notice that, in contrast to [24, Appendix A.7], we consider  $\mathbf{V}$  as a complex valued matrix, and vectors  $\mathbf{u}$  and  $\mathbf{v}$  as complex valued vectors. For such a case, expression (A.368) in [24, Appendix A.7] becomes equations (A.1) and (A.2).

The BMR solution in (29) is obtained using the next differentiation rule [39, Appendix],

$$\frac{\partial}{\partial \mathbf{W}} \text{tr}\{\mathbf{W} \mathbf{U} \mathbf{W}^+\} = 2\mathbf{W} \mathbf{U}, \quad (\text{A.4})$$

where  $\mathbf{U}$  and  $\mathbf{W}$  are arbitrary complex valued square matrices.

The derivative expressions in (38) and (39), employed to produce the assessed ML-inspired approach in (44), make use of the matrix differentiation formulas [17], [18], [24, Appendix A.7]

$$\frac{\partial}{\partial \mathbf{v}} \ln \det\{\mathbf{U}\} = \text{tr}\left\{\mathbf{U}^{-1} \frac{\partial}{\partial \mathbf{v}} \mathbf{U}\right\}, \quad (\text{A.5})$$

$$\frac{\partial}{\partial \mathbf{v}} \mathbf{U}^{-1} = -\mathbf{U}^{-1} \left( \frac{\partial}{\partial \mathbf{v}} \mathbf{U} \right) \mathbf{U}^{-1}, \quad (\text{A.6})$$

that hold for an arbitrary complex valued invertible matrix  $\mathbf{U}$ , which is function of vector  $\mathbf{v}$ ; and the formula

$$\frac{\partial}{\partial \mathbf{v}} \mathbf{D}(\mathbf{v}) = \mathbf{I}. \quad (\text{A.7})$$

Finally, the WCF-based solution in (51) considers the next differentiation rule [24, Appendix A.7],

$$\frac{\partial}{\partial \mathbf{V}} \text{tr}\{\mathbf{U} \mathbf{V} \mathbf{W}\} = \frac{\partial}{\partial \mathbf{V}^+} \text{tr}\{\mathbf{U} \mathbf{V}^+ \mathbf{W}\} = \mathbf{U}^+ \mathbf{W}^+, \quad (\text{A.8})$$

in which matrices  $\mathbf{U}$  and  $\mathbf{W}$  (not necessarily square matrices) are not function of matrix  $\mathbf{V}$ .

## REFERENCES

- [1] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data", *IEEE Trans. Geosc. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.
- [2] G. Fornaro, F. Serafino, and F. Soldovier, "Three-dimensional focusing with multipass SAR data", *IEEE Trans. Geosc. Remote Sens.*, vol. 41, no. 3, pp. 507–517, March 2003.
- [3] F. Gini, F. Lombardini and M. Montanari, "Layover solution in multibaseline SAR interferometry", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no.4, pp. 1344–1356, Oct. 2002.
- [4] M. Nannini, R. Scheiber, and A. Moreira, "Estimation of the minimum number of tracks for SAR tomography", *IEEE Trans. Geosc. Remote Sens.*, vol. 47, no. 2, pp. 531–543, Jan. 2009.
- [5] M. Nannini, A. Reigber and R. Scheiber, "A study on irregular baseline constellations in SAR tomography", in *EUSAR*, 2010, pp. 1–4.
- [6] M. Nannini, R. Scheiber, R. Horn, and A. Moreira, "First 3-D reconstructions of targets hidden beneath foliage by means of polarimetric SAR tomography", *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no.1, pp. 60–64, Jan. 2012.
- [7] Y. Huang, L. Ferro-Famil and A. Reigber, "Under-Foliage Object Imaging Using SAR Tomography and Polarimetric Spectral Estimators", *IEEE Trans. Geosc. Remote Sens.*, vol. 50, no. 6, pp. 2213–2225, June 2012.
- [8] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, vol. 1, Upper Saddle River, NJ: Prentice hall, 2005.
- [9] J. Li, P. Stoica and Z. Wang, "Robust Capon Beamforming", in *Robust Adaptive Beamforming*, Li and P. Stoica Eds, NJ: Wiley-Interscience, 2005, pp. 91–200.
- [10] E. Aguilera, M. Nannini and A. Reigber, "A Data-Adaptive Compressed Sensing Approach to Polarimetric SAR Tomography of Forested Areas", *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no.3, pp. 543–547, Sept. 2012.
- [11] E. Aguilera, M. Nannini and A. Reigber, "Wavelet-Based Compressed Sensing for SAR Tomography of Forested Areas", *IEEE Trans. Geosc. Remote Sens.*, vol. 51, no.12, pp. 5283–5295, Dec. 2013.
- [12] M. Schmitt and U. Stilla, "Maximum-likelihood-based approach for single-pass synthetic aperture radar tomography over urban areas", *IET Radar, Sonar & Navigation*, vol. 8, no. 9, pp. 1145–1153, April 2014.
- [13] G. D. Martin del Campo, M. Nannini, and A. Reigber, "Towards Feature Enhanced SAR Tomography: A Maximum-Likelihood Inspired Approach", *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 11, pp. 1730–1734, Nov. 2018.
- [14] S. Tebaldini, "Algebraic synthesis of forest scenarios from multibaseline PolInSAR data", *IEEE Trans. Geosc. Remote Sens.*, vol. 47, no. 12, pp. 4132–4142, Dec. 2009.
- [15] S. Tebaldini, "Single and multipolarimetric SAR tomography of forested areas: A parametric approach", *IEEE Trans. Geosc. Remote Sens.*, vol. 48, no. 5, pp. 2375–2387, May 2010.
- [16] S. Tebaldini and F. Rocca, "Multibaseline polarimetric SAR tomography of a boreal forest at P-and L-bands", *IEEE Trans. Geosc. Remote Sens.*, vol. 50, no. 1, pp. 232–246, Jan. 2012.
- [17] Y. V. Shkvarko, "Estimation of wavefield power distribution in the remotely sensed environment: Bayesian maximum entropy approach", *IEEE Trans. Sign. Process.* vol. 50, no. 9, pp. 2333–2346, Nov. 2002.
- [18] Y. V. Shkvarko, "Unifying regularization and Bayesian estimation methods for enhanced imaging with remotely sensed data—Part I and Part II", *IEEE Trans. Geosc. Remote Sens.*, vol. 42, no. 5, pp. 923–940, May 2004.
- [19] Y. V. Shkvarko, "Unifying experiment design and convex regularization techniques for enhanced imaging with uncertain remote sensing data—Part I and Part II", *IEEE Trans. Geosc. Remote Sens.*, vol. 48, no. 1, pp. 82–111, Jan. 2010.
- [20] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging*. CRC press, 1998.
- [21] H.H. Barrett and K.J. Myers, *Foundations of Image Science*, NY: Wiley, 2004.
- [22] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University press, 2004.
- [23] M. Fuhry and L. Reichel, "A new Tikhonov regularization method", *Numerical Algorithms*, vol. 59, no. 3, pp. 433–445, March 2012.
- [24] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. NY: John Wiley & Sons, 2004.
- [25] Y. T. Kostenko and Y. V. Shkvarko, "Combined statistical regularization and experiment-design-theory-based nonlinear techniques



for extended objects imaging from remotely sensed data”, in *SPIE's International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, International Society for Optics and Photonics, 1994, pp. 309–317.

- [26] T. Yardibi, J. Li, P. Stoica, M. Xue and A. B. Baggeroer, “Source Localization and Sensing: A Nonparametric Iterative Adaptive Approach Based on Weighted Least Squares”, *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no.1, pp. 425–443, Jan. 2010.
- [27] G. Martín del Campo, Y. Shkvarko, A. Reigber and M. Nannini, “TomoSAR imaging for the study of forested areas: A virtual adaptive Beamforming approach. *Remote Sensing*, vol. 10, no. 11, Nov. 2018.
- [28] P. Stoica, P. Babu and J. Li, “New Method of Sparse Parameter Estimation in Separable Models and Its Use for Spectral Analysis of Irregularly Sampled Data”, *IEEE Trans. Sign. Process.*, vol. 59, no. 1, pp. 35–47, Oct. 2010.
- [29] P. Stoica, P. Babu and J. Li, “SPICE: A Sparse Covariance-Based Estimation Method for Array Processing”, *IEEE Trans. Sign. Process.*, vol. 59, no. 2, pp. 629–638, Nov. 2010.
- [30] C. R. Rojas, D. Katselis and H. Hjalmarsson, “A Note on the SPICE Method”, *IEEE Trans. Sign. Process.*, vol. 61, no. 18, pp. 4545–4551, July 2013.
- [31] Y. V. Shkvarko, J. Tuxpan, and S. R. Santos, “Dynamic experiment design regularization approach to adaptive imaging with array radar/SAR sensor systems”, *Sensors*, vol. 2011, No 5, pp. 4483–4511.
- [32] S. Har-Peled and B. Raichel, “The Fréchet distance revisited and extended”, *ACM Transactions on Algorithms (TALG)*, vol. 10, no. 1, Jan. 2014.
- [33] T. Eiter and H. Mannila, “Computing discrete Fréchet distance”. Tech. Report CD-TR 94/64, Information Systems Department, Technical University of Vienna. April, 1994.
- [34] CVXPY. Available online: [www.cvxpy.org](http://www.cvxpy.org)
- [35] S. Ramani, T. Blu and M. Unser, “Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms”, *IEEE Trans. on Image Processing*, vol. 17, no. 9, pp. 1540–1554, Sept. 2008.
- [36] O. D’Hondt, C. López-Martínez, S. Guillaso and O. Hellwich, “Nonlocal Filtering Applied to 3-D Reconstruction of Tomographic SAR Data”, *IEEE Trans. Geosc. Remote Sens.*, vol. 56, no. 1, pp. 272–285, Jan. 2018.
- [37] H. Aghababae, G. Ferraioli, G. Schirinzì and M. R. Sahebi, “The role of nonlocal estimation in SAR tomographic imaging of volumetric media”, *IEEE Geosc. Remote Sens. Letters*, vol. 15, no. 5, pp. 729–733, March. 2018.
- [38] V. Cazcarra-Bes, M. Tello-Alonso, R. Fischer, M. Heym and K. Papathanassiou, “Monitoring of Forest Structure Dynamics by Means of L-Band SAR Tomography”. *Remote Sensing*, vol. 9, no.12, pp. 1229 – 1250, Nov. 2017.
- [39] Y. Shkvarko, “From matched spatial filtering towards the fused statistical descriptive regularization method for enhanced radar imaging”. *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1-9, Nov. 2005.



**Gustavo Martín del Campo** received the degree in electronics and communications engineering from the University of Guadalajara, Mexico, in 2008. In 2013 and 2017, he obtained the M. Sci. and Dr. Sci. (PhD equivalent) degrees in electrical engineering, with specialization in telecommunications, from the Centre for Research and Advanced Studies (Cinvestav) of the National Polytechnic Institute (IPN), Mexico. Since 2017 he is with the Microwaves and Radar Institute (HR) of the German Aerospace Center (DLR). His research interests are in applications of signal processing to remote sensing, particularly SAR tomography, inverse problems, random fields estimation and adaptive spatial analysis.



**Matteo Nannini** received the Laurea degree in telecommunication engineering from the University of Florence, Italy, in 2003. The thesis has been done in collaboration with the Microwaves and Radar Institute (HR) of the German Aerospace Center (DLR). In 2009 he obtained the PhD degree in electrical engineering from the Karlsruhe Institute of Technology (KIT) with a thesis on SAR tomography. Since 2004 he is with the Microwaves and Radar Institute of DLR where he is working on SAR tomography, spectral estimation techniques, radar ice sounding, and PSI time series analysis. He is involved in several studies for the evaluation of the performance for future bi- and multi-static SAR missions.



**Andreas Reigber** (M’02–SM’10–F’16) received the Diploma degree in physics from the University of Konstanz, Konstanz, Germany, in 1997, the Ph.D. degree from the University of Stuttgart, Stuttgart, Germany, in 2001, and the Habilitation degree from the Berlin University of Technology, Berlin, Germany, in 2008. He is currently the Head of the SAR Technology Department, Microwave and Radar Institute, German Aerospace Center, Weßling, Germany, where he is leading the development and operation of state-of-the-art airborne SAR sensors. He is also a Professor of remote sensing and digital image processing with the Berlin University of Technology, Berlin, Germany. His research interests include various aspects of multimodal, multichannel, and high-resolution SAR processing and post-processing. Dr. Reigber has received several prize paper awards, among them the IEEE TGRS Transactions Prize Paper Award in 2001 and 2016 for his works on polarimetric SAR tomography and nonlocal speckle filtering, respectively, and also the IEEE TGRS Letters Prize Paper Award in 2006 for his work on multi-pass SAR processing.