

# Constrained Tensor Decompositions for SAR Data: Agricultural Polarimetric Time Series Analysis

Nikita Basargin<sup>1</sup>, Alberto Alonso-González<sup>2</sup>, and Irena Hajnsek<sup>3</sup>, *Fellow, IEEE*

**Abstract**—Tensor decompositions are a powerful tool for multidimensional data analysis, interpretation, and signal processing. This work develops a constrained tensor decomposition framework for complex multidimensional synthetic aperture radar (SAR) data. The framework generalizes the canonical polyadic (CP) decomposition by formulating it as an optimization problem and allows precise control over the shape and properties of the output factors. The implementation supports complex tensors, automatic differentiation, and different loss functions and optimizers. We discuss the importance of constraints for physical validity, interpretability, and uniqueness of the decomposition results. To illustrate the framework, we formulate a polarimetric time series decomposition and apply it to data acquired over agricultural areas to analyze the development of four crop types at the X-, C-, and L-bands over the period of 12 weeks. The obtained temporal factors describe the changes in the crops in a compact way and show a correlation to certain crop parameters. We extend the existing polarimetric time series change analysis with the decomposition to show the changes in more detail and provide an interpretation through the polarimetric factors. The decomposition framework is extensible and promising for joint information extraction from multidimensional SAR data.

**Index Terms**—Agriculture, constraints, optimization, polarimetric synthetic aperture radar (PolSAR), SAR, tensor decomposition, time series.

## I. INTRODUCTION

SYNTHETIC aperture radar (SAR) is widely used for remote sensing and Earth observation with a large variety of applications in different research areas [1]. In contrast to optical sensors, SAR has a sensitivity to dielectric and geometrical properties and offers unique advantages such as cloud penetration. SAR data are defined in the complex domain where both amplitude and phase carry information. Depending on the sensor and the acquisition setup, the data have different

dimensions: SAR images naturally offer the spatial dimension, polarimetry adds information about wave polarization, interferometry and tomography combine the phase information from two or more acquisitions to obtain sensitivity to the vertical scattering profile, temporal image stacks capture changes over time, and radar frequency bands offer sensitivity to different object properties. Analyzing and interpreting this high-dimensional and complex data is often challenging and requires advanced processing methods.

SAR decompositions have been commonly used to reduce the data to a few key parameters and simplify the analysis. For polarimetry, many different decompositions have been proposed including Cloude and Pottier [2], Freeman and Durden [3], Krogager [4], and Yamaguchi [5]. A detailed overview can be found in [6]. The parameters obtained from the decomposition are helpful for unsupervised classification and land cover characterization. More recently, decompositions for two data dimensions have been proposed. One example is the sum of Kronecker products (SKP) decomposition [7] for polarimetric multibaseline data that jointly uses information from polarimetry and interferometry/tomography.

For data that can be represented as a matrix, several algebraic matrix factorization and decomposition methods are available including eigendecomposition, singular value decomposition (SVD), principle compound analysis (PCA), and nonnegative matrix factorization (NMF) [8], [9], [10], [11]. Some SAR decompositions build on top of the algebraic decompositions, for example, eigendecomposition and SVD are used in Cloude decomposition and SKP decomposition, respectively.

Combining multiple data dimensions is a promising approach to extract additional information or improve prediction accuracy. Matrix factorization methods can be applied to multidimensional data tensors but require reshaping tensors into matrices before factorization. One example is stacks of 2-D images or matrices. Before factorization, each image is reshaped into a vector that then forms columns of a new matrix to be factorized. Since some data properties are not preserved during reshaping, additional steps might be required before component interpretation, as in the SKP decomposition.

Tensor decompositions generalize the matrix decompositions and directly work with multidimensional tensors without the need to reshape them into matrices. Several tensor decompositions exist, the most common being the canonical polyadic (CP) (also known as CANDECOMP or PARAFAC) [12], [13] and Tucker [14] decompositions. They are powerful tools to jointly analyze multidimensional data and are used in several

Manuscript received 10 March 2023; revised 13 October 2023; accepted 5 November 2023. Date of publication 8 November 2023; date of current version 29 November 2023. The work of Nikita Basargin was supported by the Helmholtz Association under the joint research school Munich School for Data Science (MUDS). (*Corresponding author: Nikita Basargin.*)

Nikita Basargin is with the Microwaves and Radar Institute, German Aerospace Center (DLR), 82234 Wessling, Germany, and also with the TUM School of Life Sciences, Technical University of Munich (TUM), 85354 Freising, Germany (e-mail: nikita.basargin@dlr.de).

Alberto Alonso-González was with the Microwaves and Radar Institute, German Aerospace Center (DLR), 82234 Wessling, Germany. He is now with the Signal Theory and Communications Department, Technical University of Catalonia (UPC), 08034 Barcelona, Spain (e-mail: alberto.alonso-gonzalez@upc.edu).

Irena Hajnsek is with the Microwaves and Radar Institute, German Aerospace Center (DLR), 82234 Wessling, Germany, and also with the Institute of Environmental Engineering, ETH Zürich, 8049 Zürich, Switzerland (e-mail: irena.hajnsek@dlr.de).

Digital Object Identifier 10.1109/TGRS.2023.3331599

research areas. The applications include multidimensional data analysis, compression, classification, dimensionality reduction, and many others [15], [16]. A complete overview of tensor decompositions is given in [17].

In the field of remote sensing, tensor decompositions have been widely applied to optical hyperspectral data [18]. However, complex-valued SAR data are inherently different from optical images both in terms of structure and dynamic range. Therefore, the existing tensor decomposition methods often require modifications to be applied to SAR. Recently, different methods involving tensor decompositions have specifically been developed for SAR data. In the context of image classification, polarimetric features have first been extracted before applying Tucker decomposition in [19]. Tensor decompositions have also been used for interferometric SAR data stack filtering and urban mapping in [20]. A despeckling method for polarimetric time series was proposed in [21].

Tensor decompositions are promising tools for multidimensional SAR data processing and analysis. In this article, we discuss the CP decomposition, its limitations when applied to SAR data, and propose some solutions to resolve these limitations. The first limitation is that the standard formulation of CP produces vector factors, while some dimensions such as polarimetry typically represent the data in matrices. The second limitation is that not all the decomposition solutions may be physically valid. Therefore, it is important to constrain the decomposition solution space and enforce a specific structure on the resulting factors. For example, polarimetric coherency matrices have to be Hermitian and positive semidefinite (PSD).

This article introduces a constrained tensor decomposition framework for SAR data that addresses these limitations. In Section II, we discuss the standard CP decomposition and extend it into a framework that allows to define constraints on factors and allows arbitrary factor shapes. To demonstrate the framework, we formulate a specific decomposition and use it to analyze polarimetric time series which is an active research area with recent works by Alonso-González et al. [22] and Cloude [23]. To visualize the results, we extend the change analysis from [22] with the decomposition factors. In Section III, we evaluate the decomposition in the context of crop monitoring and change detection. We highlight the effect of constraints on solution uniqueness, analyze changes for different crop types and bands, and evaluate the correlation between the decomposition factors and selected crop parameters. In Section IV, we discuss the findings and present possible extensions to the framework. Section V concludes this article.

## II. METHODS

### A. Notation

The notation mainly follows the definitions used by Kolda and Bader [17]. We denote vectors with bold lowercase letters, e.g.,  $\mathbf{a}$ , matrices with bold capital letters, e.g.,  $\mathbf{M}$ , and tensors with capital calligraphic letters, e.g.,  $\mathcal{X}$ . We denote the individual elements of tensor  $\mathcal{X}$  found at index  $[i_1, \dots, i_m]$  by  $\mathcal{X}_{[i_1, \dots, i_m]}$ .

### B. SAR Polarimetry

A fully polarimetric SAR system provides the scattering matrix  $\mathbf{S}$  for each pixel with matrix elements representing the complex scattering coefficients of the received and transmitted horizontal (H) and vertical (V) polarization combinations

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix}. \quad (1)$$

While the matrix is able to completely describe deterministic scatterers, a second-order formalism is required for distributed scatterers [1]. The elements can be arranged into the scattering vector  $\mathbf{k}$  in the Pauli basis [24] that follows a zero-mean complex Gaussian distribution given a large number of distributed scatterers within the resolution cell:

$$\mathbf{k} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ S_{HV} + S_{VH} \end{bmatrix}. \quad (2)$$

The covariance matrix  $\mathbf{T}$  describing the Gaussian distribution (also known as coherency matrix) can be estimated by combining vectors  $\mathbf{k}_i$  from  $n$  pixels

$$\mathbf{T} = \frac{1}{n} \sum_{i=1}^n \mathbf{k}_i \circ \mathbf{k}_i^* \quad (3)$$

where  $*$  denotes the complex conjugation, and  $\circ$  represents the outer product.  $\mathbf{T}$  completely describes the polarimetric response of distributed scatterers under the Gaussian hypothesis and it is commonly used for polarimetric SAR analysis. The matrix  $\mathbf{T}$  is Hermitian, PSD, and has real and nonnegative eigenvalues by construction.

The choice of the Pauli basis has an advantage when it comes to interpreting the scattering mechanisms [25]. We refer to the diagonal elements of  $\mathbf{T}$  as Pauli elements since they correspond to the backscattered power of the Pauli components in the  $\mathbf{k}$  vector. The first element  $\mathbf{T}_{[1,1]}$  (HH + VV polarization) can be associated with surface scattering, the second  $\mathbf{T}_{[2,2]}$  (HH - VV) indicates dihedral scattering and the third  $\mathbf{T}_{[3,3]}$  (HV + VH) can be related to volume scattering. A common Pauli RGB color scheme displays the first, second, and third elements in blue, red, and green, respectively.

### C. CP Tensor Decomposition

The CP decomposition factorizes a tensor into a sum of components. Each component is constructed from an outer product of factor vectors, with one factor for each dimension. In the 3-D case, a tensor  $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$  decomposes into  $R$  components each defined by the factor vectors  $\mathbf{a}_r \in \mathbb{C}^I$ ,  $\mathbf{b}_r \in \mathbb{C}^J$ , and  $\mathbf{c}_r \in \mathbb{C}^K$

$$\mathcal{X} \approx \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r. \quad (4)$$

The reconstruction function  $r(\cdot)$  reconstructs the tensor  $\mathcal{R}$  from the tuple  $F$  that contains all the factors

$$r(F) = \mathcal{R} \quad (5)$$

with  $F = (\mathbf{a}_1, \dots, \mathbf{a}_R, \mathbf{b}_1, \dots, \mathbf{b}_R, \mathbf{c}_1, \dots, \mathbf{c}_R)$ . Any tensor  $\mathcal{X}$  can be perfectly reconstructed ( $\mathcal{X} = \mathcal{R}$ ) by the

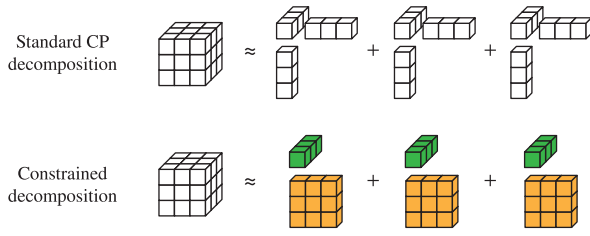


Fig. 1. Standard CP decomposition and the proposed decomposition with constraints and flexible factor shapes. The colors indicate different constraints applied to the factors.

CP decomposition given a sufficiently large  $R$ . The tensor rank is the smallest  $R$  that fully reconstructs the tensor. Each of the decomposition components is a rank-1 tensor.

There exist several algorithms performing the CP decomposition. The alternating least squares (ALS) algorithms optimize one factor dimension at a time, keeping the others fixed and repeating the process for all the dimensions until convergence. Another class of algorithms directly obtains the factors using gradient-based optimization. An extensive overview of the CP decomposition and the decomposition algorithms can be found in [17] and [26].

#### D. Constrained Tensor Decomposition Framework

Here, we propose a decomposition framework that generalizes the standard CP decomposition with two main extensions: allow for flexible factor shapes and introduce different factor constraints, as illustrated in Fig. 1.

Flexible factor shapes remove the restriction for factors to be vectors and allow them to be matrices or even other tensors. This is especially useful when working with polarimetric SAR data that are commonly expressed through covariance matrices.

Factor constraints restrict the solution space for the decomposition output. In the context of SAR, constraints are used to eliminate solutions that are physically nonvalid, such as negative backscatter powers or covariance matrices that are not PSD.

We formulate the decomposition as an optimization problem and use the solution of the problem to obtain the constrained factors. The differentiable function to be optimized consists of three main steps: applying the factor constraints, reconstructing an approximation of the original tensor, and computing the loss, as shown in Fig. 2.

First, the constraint function  $c(\cdot)$  maps the tuple of unconstrained factors  $F_u$  to the tuple of constrained factors  $F_c$

$$F_c = c(F_u). \quad (6)$$

Then, the reconstruction function  $r(\cdot)$  transforms the constrained factors into the reconstructed tensor  $\mathcal{R}$

$$\mathcal{R} = r(F_c). \quad (7)$$

Finally, the loss function  $l(\cdot)$  evaluates the difference between the tensors  $\mathcal{X}$  and  $\mathcal{R}$

$$L = l(\mathcal{X}, \mathcal{R}). \quad (8)$$

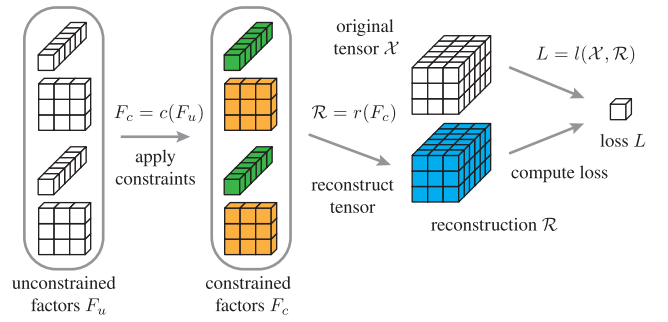


Fig. 2. Formulation of the constrained decomposition as an optimization problem. The function to be optimized consists of three main parts: constraining factors, reconstruction of the approximation tensor, and loss calculation.

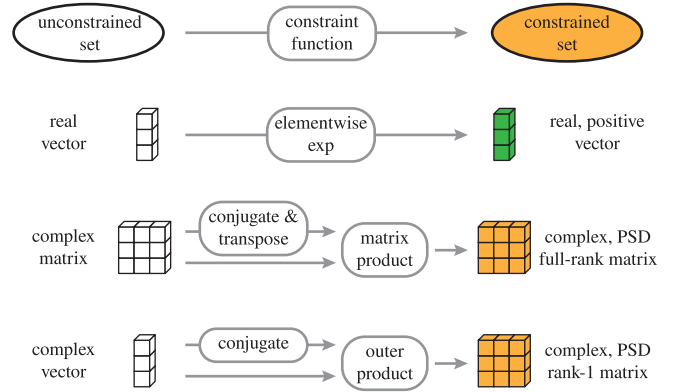


Fig. 3. Constraint implementation: the set of unconstrained factors is mapped to a constrained factor set. Examples for positive vectors ( $\mathbb{R}^+$ ) and PSD matrices (both full-rank and rank-1).

1) *Constraint Implementation:* The factor constraints are realized by mapping the unconstrained set onto the constrained set. Examples of constraints illustrate the idea in Fig. 3. To achieve a factor with only positive real values, we start from a vector with any real numbers and apply the exponential function elementwise. A more advanced example is the PSD constraint on matrices. One way to implement it is to start with an arbitrary full-rank matrix and multiply it by the conjugate transpose of itself, resulting in a constrained full-rank PSD matrix. Another way is to start with a vector and form an outer product between the vector and the conjugate of itself, resulting in a rank-1 PSD matrix. More details on constrained optimization on manifolds and a list of constraints can be found in [27].

We apply factor constraints to each factor dimension individually and allow different dimensions to have different constraints. We denote the function that maps the tuple of all the unconstrained factors to the tuple of all the constrained factors with  $c(\cdot)$  and the dimension-specific functions with  $c^{\dim}(\cdot)$  where  $\dim$  is the dimension.

2) *Reconstruction Function:* The reconstruction function  $r(\cdot)$  transforms the constrained factors to the reconstructed tensor. In the standard CP decomposition, the reconstruction is a sum of outer products of factors as indicated in (4). In the constrained decomposition framework, we use the same reconstruction function. However, the factors are no longer limited to vectors and can be matrices or tensors.

The outer product of two tensors  $\mathcal{U} \in \mathbb{C}^{I_1 \times \dots \times I_M}$  and  $\mathcal{V} \in \mathbb{C}^{J_1 \times \dots \times J_N}$  is a tensor  $\mathcal{W} \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$ . For example, a product of a  $3 \times 42$  matrix and a seven-element vector creates a  $3 \times 42 \times 7$  tensor. Similar to the vector case, the entries of  $\mathcal{W}$  are products of combinations of the entries in  $\mathcal{U}$  and  $\mathcal{V}$

$$\mathcal{W}_{[i_1, \dots, i_m, j_1, \dots, j_n]} = \mathcal{U}_{[i_1, \dots, i_m]} \mathcal{V}_{[j_1, \dots, j_n]}. \quad (9)$$

After forming the outer products of factors, the resulting component tensors are summed up to obtain the final reconstruction  $\mathcal{R}$ .

3) *Loss Function*: The loss function  $l(\cdot)$  measures the difference between two input tensors and outputs a single scalar  $L \in \mathbb{R}$ . Let  $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_N}$ , then the norm of  $\mathcal{X}$  is

$$\|\mathcal{X}\| = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_N=1}^{I_N} |\mathcal{X}_{[i_1, i_2, \dots, i_N]}|^2}. \quad (10)$$

The standard CP decomposition uses the squared error loss function which can be expressed using the tensor norm

$$l(\mathcal{X}, \mathcal{R}) = \|\mathcal{X} - \mathcal{R}\|^2. \quad (11)$$

Some formulations use the mean squared error resulting in a slightly different loss value; however, it has no effect on the optimization solution.

The choice of the loss function implicitly assumes the distribution of the data. For example, the squared error loss assumes a Gaussian distribution. A recent CP formulation in [26] allows us to flexibly choose other loss functions and lists several candidates for different data distributions. Our proposed framework shares this flexibility allowing to use any differentiable function mapping  $\mathcal{X}$  and  $\mathcal{R}$  to  $L$ . In this article, we use the standard squared error loss and discuss optional regularization in Section III-B2.

### E. Decomposition With Optimization

Function optimization is a large research area with well-developed tools, formalism, and implementations. Our goal is to minimize the loss  $L$  and obtain the factors best reconstructing the tensor  $\mathcal{X}$

$$L = l(\mathcal{X}, \mathcal{R}) = l(\mathcal{X}, r(F_c)) = l(\mathcal{X}, r(c(F_u))). \quad (12)$$

When the function gradient is known, efficient minimization is possible using first-order iterative gradient descent-based algorithms such as Adam [28]. If the number of function input parameters is relatively small and the second derivative (Hessian matrix) is available, second-order optimization methods such as Broyden–Fletcher–Goldfarb–Shanno (BFGS) [29], [30] can provide faster convergence. For the scope of this article, the specific choice of the optimizer is not of great importance. Therefore, we choose the commonly used first-order Adam optimizer.

1) *Automatic Differentiation*: To perform the optimization, we need to obtain partial derivatives of each element of each unconstrained factor in  $F_u$  with respect to the loss  $L$ . An analytical expression can be derived by hand applying the chain rule. However, when experimenting with the decomposition framework, adapting constraints, and adding dimensions,

having to derive an analytical solution after every change can be time-consuming. Therefore, we suggest using automatic differentiation to compute the gradient in an efficient way when the function is evaluated.

We use PyTorch [31] to perform optimization and automatic differentiation for the decomposition. PyTorch is a widely adopted machine learning framework with a large community, which provides an efficient implementation of many tensor operations, enables hardware acceleration, and supports complex tensors and differentiation through Wirtinger calculus. Other frameworks such as TensorFlow [32], JAX [33], or Keras [34] can also be used for automatic differentiation. The degree of support for complex differentiation varies but is improving over time.

2) *Decomposition Factor Estimation*: To find the constrained decomposition factors, we start with randomly initialized unconstrained factors and iteratively perform optimization steps until convergence. During each optimization step, we first apply the constraints to obtain the constrained factors. Then, we reconstruct  $\mathcal{R}$  using the reconstruction function and compute the loss  $L$ . We obtain the gradients with respect to the unconstrained factors through automatic differentiation and use them to update the unconstrained factors. The optimization steps are repeated until convergence. Finally, we apply the constraints to the unconstrained factors and obtain the constrained decomposition output.

### F. Application: Polarimetric Time Series Decomposition

In this section, we formulate a specific decomposition problem to illustrate the proposed decomposition framework. We decompose fully polarimetric time series data into  $R$  components made up of 1-D temporal and 2-D polarimetric factors. The decomposition only uses temporal and polarimetric information keeping the number of dimensions small for straightforward interpretation. At the same time, decomposing a 3-D tensor is sufficient to illustrate the contributions of the proposed framework including different factor shapes and constraints. The decomposition assumes  $R$  underlying mechanisms responsible for the backscatter signal. These mechanisms keep the same polarimetric signature but change their intensity as the scene changes over time.

We decompose the data tensor  $\mathcal{X} \in \mathbb{C}^{N \times 3 \times 3}$  created from a stack of  $N$  polarimetric  $3 \times 3$  covariance matrices. Each of the  $R$  decomposition components is defined by a polarimetric  $3 \times 3$  matrix factor  $\mathbf{P}_r$  describing the scattering mechanism and an  $N$ -element temporal factor vector  $\mathbf{t}_r$  describing how the intensity of the component changes over time. The detailed compute graph is shown in Fig. 4.

To ensure the physical validity and interpretability of the decomposition result, we constrain  $\mathbf{P}_r$  to be a PSD matrix and  $\mathbf{t}_r$  to only contain real and positive values. The tuples  $F_u$  and  $F_c$  of unconstrained and constrained factors are given by

$$F_u = (\mathbf{t}_{u,1}, \dots, \mathbf{t}_{u,R}, \mathbf{p}_{u,1}, \dots, \mathbf{p}_{u,R}) \quad (13)$$

$$F_c = (\mathbf{t}_1, \dots, \mathbf{t}_R, \mathbf{P}_1, \dots, \mathbf{P}_R) \quad (14)$$

where  $\mathbf{t}_{u,r} \in \mathbb{R}^N$  and  $\mathbf{p}_{u,r} \in \mathbb{C}^3$  are the unconstrained temporal and polarimetric factors, respectively. The functions

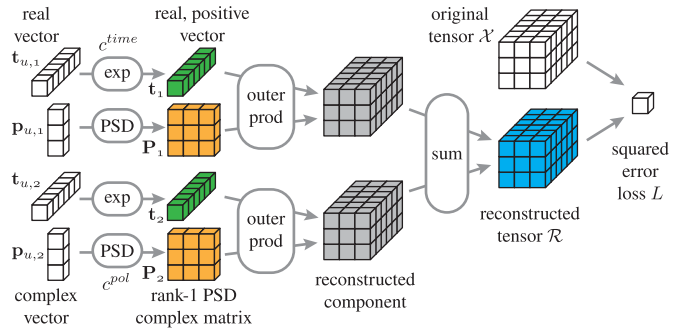


Fig. 4. Compute graph of polarimetric time series decomposition. Two components are shown in this example.

constraining the individual factor dimensions are given by

$$\mathbf{t}_r = c^{\text{time}}(\mathbf{t}_{u,r}) = \exp(\mathbf{t}_{u,r}) \quad (15)$$

$$\mathbf{P}_r = c^{\text{pol}}(\mathbf{p}_{u,r}) = \mathbf{p}_{u,r} \circ \mathbf{p}_{u,r}^* \quad (16)$$

where  $\exp$  denotes the exponential function applied elementwise. Constructing  $\mathbf{P}_r$  from an outer product of vectors has the additional effect of constraining the matrix to be rank-1. It is also possible to use full-rank polarimetric factors for each component of the decomposition and we discuss the differences in Section III-B.

The reconstruction function is a sum of the outer products of the constrained vector and matrix factors

$$\mathcal{X} \approx \mathcal{R} = \sum_{r=1}^R \mathbf{t}_r \circ \mathbf{P}_r. \quad (17)$$

To avoid scaling ambiguities, we normalize the temporal factor elements to sum up to one

$$\sum_{n=1}^N \mathbf{t}_{r[n]} = 1. \quad (18)$$

The scaling magnitude is then absorbed into the corresponding polarimetric factor  $\mathbf{P}_r$ . Temporal factor normalization highlights the relative changes and makes it easier to compare how different components change over time. After normalization, the weight  $w_r$  of the component can be obtained as the trace of  $\mathbf{P}_r$

$$w_r = \text{Tr} \mathbf{P}_r. \quad (19)$$

The relative weight of the individual component is defined by

$$w_r^{\text{rel}} = \frac{w_r}{\sum_{i=1}^R w_i}. \quad (20)$$

We order the components by their weights in descending order, so that the first component has the highest importance in the reconstruction.

As opposed to some matrix decompositions such as the SVD, the obtained factors are generally not orthogonal. Therefore, each polarimetric factor is able to represent combinations of different scattering mechanisms (e.g., dihedral and volume scattering) and is not restricted by other factors. The decomposition finds a nonorthogonal polarimetric basis that best approximates the time series data under the given constraints. The temporal factors describe the evolution of the

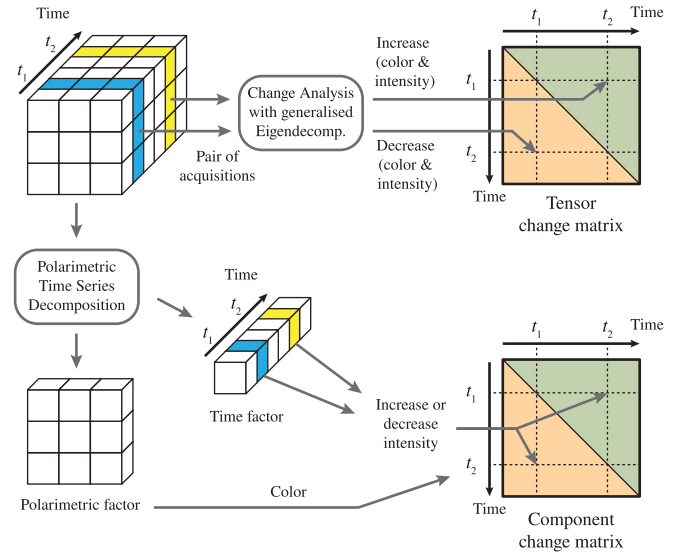


Fig. 5. Change matrices visualize increasing and decreasing signal components for each pair of acquisitions in the upper and lower triangular parts, respectively. The Pauli RGB color in the matrices is interpolated between the time points and indicates the polarization.

corresponding polarimetric factor over time. Nonorthogonality is also important for the temporal factors since orthogonal factors would be too restrictive in combination with the nonnegative constraint. Compared with a time series of an individual coherency matrix element, temporal factors are more robust to noise in the data, since they have a full polarimetric matrix attached to them.

An important difference to the existing SAR decompositions [3], [4], [5], [7] is that the number of the components can be chosen depending on the application. For example, performing the decomposition with only one component will extract a single scattering mechanism with the highest importance over the whole time frame. Using several components refines the reconstruction and allows the analysis of fine-grained changes.

Another notable property is that there are no initial assumptions on the specific scattering mechanisms. The decomposition finds the mechanisms that best describe the measured signal in a completely automated and data-driven fashion. We then use the polarimetric factors to interpret the mechanisms.

It should be noted that an alternative vector representation of second-order polarimetric information as suggested in [35] can also be used in the proposed decomposition framework. In this case, the constraint function  $c^{\text{pol}}(\cdot)$  needs to be adjusted to map unconstrained vectors to the subspace of physically feasible constrained vectors.

### G. Polarimetric Change Analysis and Visualization

To visualize and compare the decomposition results in a compact and concise manner, we use the polarimetric change matrix visualization first introduced in [22]. The method quantifies the changes in the signal for each acquisition pair and for each polarization using the generalized eigendecomposition

$$\mathbf{T}_2 \mathbf{w}_i = \lambda_i \mathbf{T}_1 \mathbf{w}_i \quad (21)$$

where  $\mathbf{T}_1$  and  $\mathbf{T}_2$  are the coherency matrices from two acquisitions, and  $\mathbf{w}_i$  and  $\lambda_i$  ( $i = 1, \dots, 3$ ) represent the generalized eigenvectors and eigenvalues, respectively.

The eigenvalues and the corresponding eigenvectors are then separated into the increasing ( $\lambda_i > 1$ ) and decreasing ( $\lambda_i < 1$ ) sets and define the color and the intensity of the changes in the visualization. The color indicates the changing polarization in the Pauli basis (see polarimetric definitions in Section II-B). HH + VV, HH - VV, and HV polarizations are displayed in blue, red, and green, respectively. The intensity depends on the amount of increase or decrease reflecting the magnitude of the generalized eigenvalues: black means no change and bright colors indicate strong changes. In the visualization, the increasing and decreasing signal components are shown in the upper and lower triangular parts of the change matrix, respectively, as illustrated in Fig. 5 (top). Details on the polarimetric change analysis and an in-depth explanation of the visualization can be found in [22].

We extend the change matrix visualization to the individual decomposition components. Within the component, the polarimetric signature is defined by the polarimetric factor and the changes in time only represent a different scale defined by the temporal factor. Therefore, the color is extracted directly from the three diagonal Pauli elements of the polarimetric factor, and the intensity of the change is defined by the weights within the temporal factor as illustrated in Fig. 5 (bottom).

### III. EXPERIMENTAL RESULTS

#### A. Dataset

The dataset to be analyzed was obtained by the DLR F-SAR [36] airborne radar during the CROPEX campaign in 2014. The imaged area is located in southern Germany near Wallerfing and covers several agricultural fields with different crop types. A region of the imaged area is shown in Fig. 6. In this article, we focus on corn, wheat, barley, and rapeseed. The time period spans about three months from mid-May to early August and covers several stages in crop development as shown in Fig. 7. Ground measurements of the crop parameters such as vegetation height, plant water content, soil moisture, and the phenological stage are available.

We perform the decomposition on individual fields where only one crop type is present. To represent the changes over the whole field with a single change matrix, the polarimetric data are averaged over all field pixels resulting in one polarimetric coherency matrix per acquisition. The coherency matrices are then stacked along the temporal dimension resulting in a 3-D tensor.

#### B. Solution Uniqueness

The constrained decomposition framework obtains the factors by starting with randomly initialized values and following the gradient to minimize the loss until convergence. It is important to keep in mind that the gradient-based minimization finds a local minimum. If the optimized function is not convex, there can be several local minima yielding different decomposition solutions.

The solution uniqueness is affected by the number of decomposition components, the applied constraints, the regularization, and the rank of the decomposed tensor. To illustrate this, we decompose a polarimetric time series obtained at the C-band over a wheat field and evaluate the convergence. We take advantage of the decomposition framework's flexibility allowing us to apply regularization and experiment with different constraints. Three different scenarios are shown in Fig. 8: rank-1 polarimetry, full-rank polarimetry, and full-rank polarimetry with regularization.

When discussing the decomposition results, we plot the change matrices for the original and the reconstructed tensors and provide the relative reconstruction error

$$\text{error}(\mathcal{X}, \mathcal{R}) = \frac{\|\mathcal{X} - \mathcal{R}\|}{\|\mathcal{X}\|}. \quad (22)$$

In addition, we plot one change matrix per decomposition component. The components are sorted by the component weights  $w_r^{\text{rel}}$  in descending order, with the most important component being the first. In every row, the first change matrix is directly derived from the original data  $\mathcal{X}$  as proposed in [22]. The second change matrix is derived from the decomposition reconstruction  $\mathcal{R}$ . Both appear very similar when the reconstruction error is low. The last three change matrices are created by performing the decomposition and then applying the visualization from [22] to each component. The visualizations also include normalized temporal factor weights for each component in a line plot.

1) *Constraints*: Constraints play an important role in obtaining a unique solution by removing excessive degrees of freedom. Fig. 8(a) shows two decompositions runs where the polarimetric factors are constrained to be rank-1 PSD matrices according to (16). Even though the factors are initialized differently, the decomposition converges to the same solution. Fig. 8(b) shows two runs with more relaxed constraints. The polarimetric factors are full-rank PSD matrices constructed from an unconstrained full-rank matrix

$$\mathbf{P}_r = \mathbf{P}_{u,r} \mathbf{P}_{u,r}^{T*}. \quad (23)$$

While the reconstruction is equal (up to numerical errors) across two runs, the factors show differences. In this case, the interpretation of the factors is ambiguous as they change from run to run.

2) *Regularization*: Regularization is a common concept to reduce the model capacity and prevent overfitting in machine learning [37]. Some of the regularization strategies such as the L2 regularization add an additional term to the loss function (see Section II-D3) that depends on the input parameters. For the polarimetric time series decomposition, after normalizing the temporal factors, we can add L2 regularization to the constrained polarimetric factors

$$l_2(\mathcal{X}, \mathcal{R}, \mathbf{P}_1, \dots, \mathbf{P}_R) = \|\mathcal{X} - \mathcal{R}\|^2 + \lambda_{l_2} \sum_{r=1}^R \|\mathbf{P}_r\|^2 \quad (24)$$

where  $\lambda_{l_2}$  is the regularization strength that balances the two objectives of high approximation accuracy and small component norm. L2 regularization penalizes large entries in the factors and guides the decomposition toward a solution

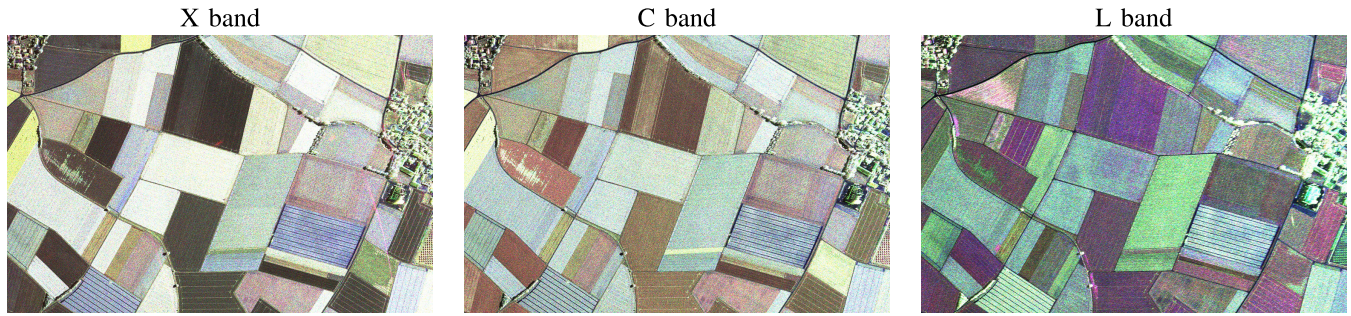


Fig. 6. Pauli images at the X-, C-, and L-bands taken on June 18 during the CROPEX 2014 campaign. Color coding: HH + VV polarization in blue, HH - VV in red, and HV + VH in green.

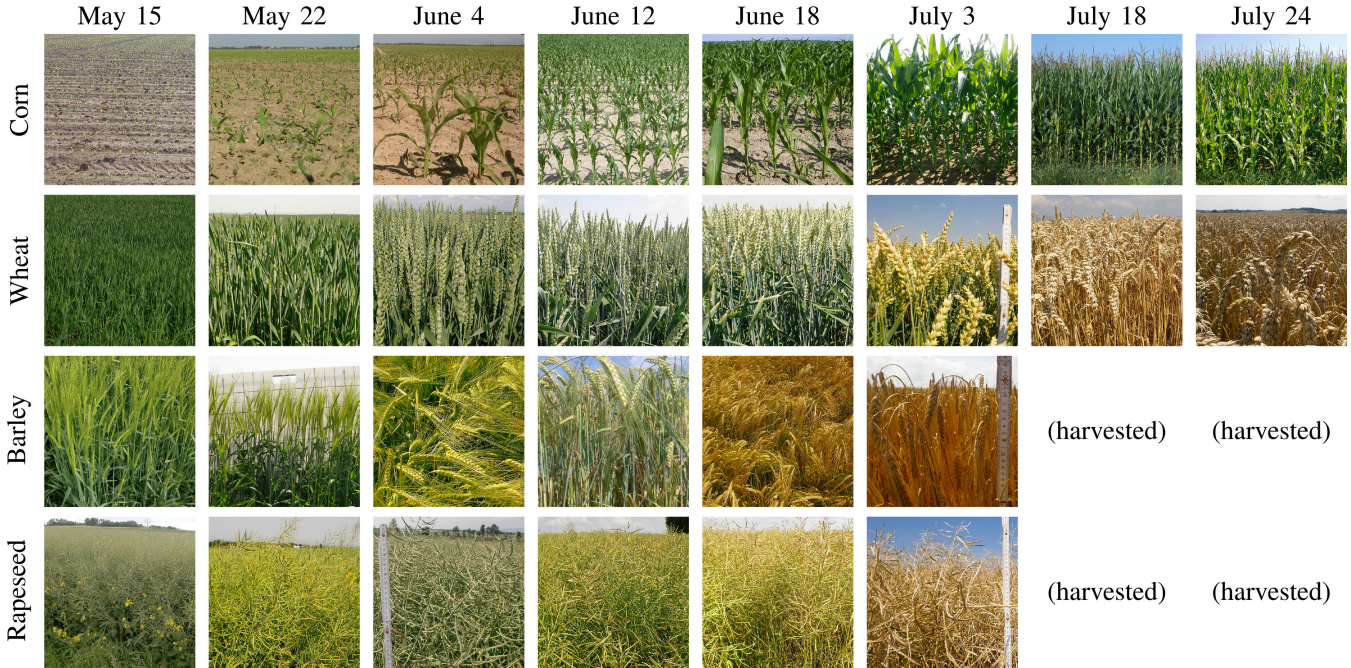


Fig. 7. Crop development during the CROPEX 2014 campaign. Corn growth was fully covered during the campaign. Barley, rapeseed, and wheat have mostly reached the maximal height at the beginning and started to dry out during the campaign.

with components having a smaller norm. From the physical perspective, the decomposition finds an approximation of the original tensor trying to use as little energy as possible.

In the context of the constrained tensor decomposition, regularization can be seen as an additional soft constrain affecting the loss landscape. When the initial choice of the constraints does not provide a unique solution [example in Fig. 8(b)], regularization can be used to guide the decomposition to a specific solution [see Fig. 8(c)]. Different regularization functions are possible each favoring specific solutions. For example, L1 regularization is known to encourage sparse solutions [26]. Note that the decomposition framework allows regularizing both the constrained and unconstrained factors. Regularizing constrained factors is generally preferred as it simplifies the physical interpretation of the regularization.

The appropriate value for the regularization strength  $\lambda_2$  can be chosen empirically depending on the application. In Sections III-C and III-D no regularization was used ( $\lambda_2 = 0$ ) as rank-1 polarimetry constraints were sufficient to obtain a unique solution.

3) *Number of Components*: The number of decomposition components directly affects the reconstruction error and the solution uniqueness. In general, the number of components should be chosen so that the reconstruction captures the original data well but filters out noise in the measurements. It should be noted that if the number of components is too high, the decomposition solution is no longer unique. This is consistent with the standard formulation of the CP decomposition. For example, infinitely many solutions are possible when decomposing a rank-1 tensor into two CP components.

If there are too few components, some parts of the data are no longer reconstructed well resulting in a large reconstruction error and a high loss value. Therefore, the loss can be used as an indicator to pick an appropriate number of components.

In this article, we use three components to obtain an accurate reconstruction and capture the changes in the data. Using more than three components does not significantly improve the reconstruction (see Fig. 9) and leads to ambiguous solutions unless regularization is used.

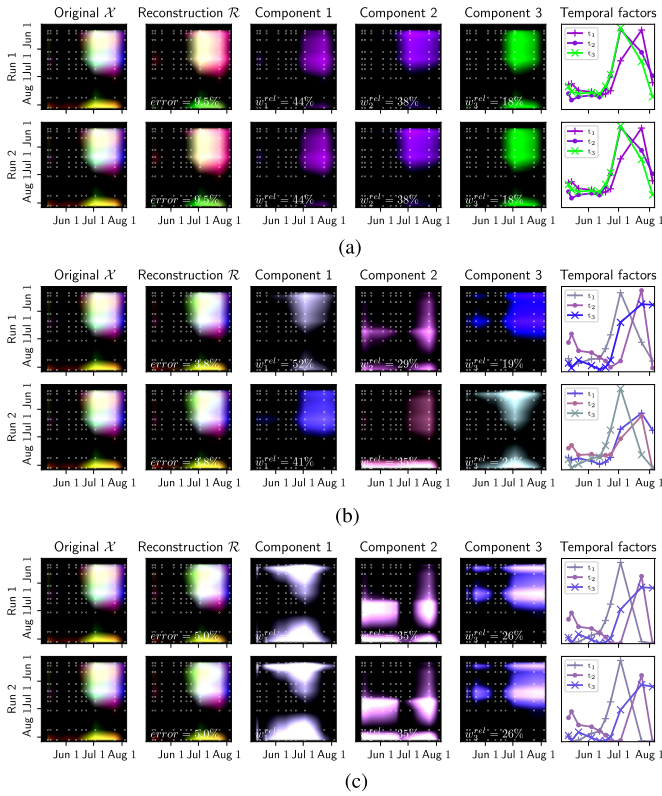


Fig. 8. Unique decomposition solution can be obtained by either applying stricter constraints (rank-1 polarimetry) or adding regularization. The same tensor (wheat at C-band) is decomposed into three components in all runs. (a) Unique solution with rank-1 PSD polarimetry (stricter constraints), no regularization. (b) Different solutions with full-rank PSD polarimetry (relaxed constraints), no regularization. (c) Unique solution with full-rank PSD polarimetry (relaxed constraints), regularization applied.

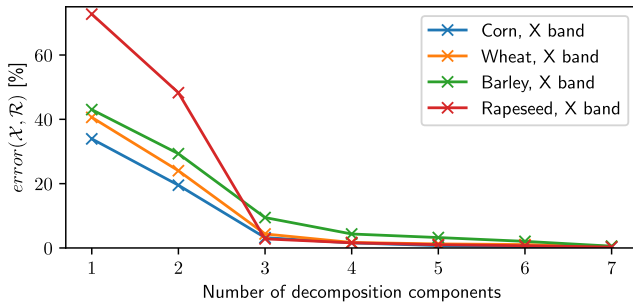


Fig. 9. Relative reconstruction error [see (22)] depending on the number of components for different crop types at the X-band.

### C. Crop Analysis

SAR is sensitive to different vegetation and soil parameters that are of interest for agricultural crop monitoring. The backscatter signal changes over time and is influenced by plant growth, changes in plant water content, changes in soil moisture, and changes to the geometrical structure of the crop.

In this section, we use the polarimetric time series decomposition in combination with polarimetric change analysis from [22] to analyze four crop types at the X-, C-, and L-bands. We decompose the measured signal into three components each having a specific polarimetric signature that changes its intensity over time. We apply rank-1 polarimetry constraints and do not use regularization.

The change matrices show distinct signatures for each crop type in different bands as shown in Fig. 10. Several processes such as plant growth, fruit maturation, drying, and harvest can be identified from the change matrices and interpreted in combination with information from ground measurements. Radar frequency bands interact differently with the crops and are sensitive to different properties and changes. On the F-SAR system [36], X-band has the shortest wavelength of 3 cm interacting with smaller objects such as leaves and having less penetration into the vegetation. On the other hand, the L-band has a longer wavelength of 23 cm, penetrating deeper into the vegetation and being more sensitive to larger scatterers and the ground. The C-band is located between the X- and L-bands with a wavelength of 6 cm.

The decomposition provides components that best describe the observations. Each of the components has a polarimetric factor defining the scattering mechanism and a temporal factor capturing the intensity changes in the scattering mechanism over time. Assuming that the changes in the signal are mainly caused by the changes in the crops, we expect to see a correlation between certain temporal factors and the crop parameters. Fig. 11 shows several cases of measured crop parameters and the temporal factors where large similarities are visible. In the following, we discuss the decomposition results and the visible changes for each crop type and band.

1) *Corn*: Corn [see Fig. 10(a)] growth period was completely covered during the campaign starting with an almost bare field and ending with vegetation reaching 3 m in height. The X- and C-bands show similar changes. There is a strong decrease in surface scattering (first component, blue) after the first two acquisitions and an increase in double bounce (red) and volume (green) components, which can be attributed to the growth of the plants, larger stems, and more leaves.

At the L-band, no decrease is observed across all the components, except for the first dates. The first component can be attributed to surface scattering based on its polarimetric signature. As the plants grow, we expect to see a stronger contribution of the corn stalks and leaves. This is clearly seen in the second and third components that represent a combination of dihedral and volume scattering and show a high correlation to the measured plant height as shown in Fig. 11(a).

The decomposition components show additional details not visible in the change matrix of the original data. For example, the X- and C-bands show a small increase in the first component (blue, surface scattering) in the middle of the growing period. This change is rather weak in relative terms and gets overpowered by larger changes in other components. However, the first component has a larger weight [indicated by  $w_1^{\text{rel}}$  in Fig. 10(a)], and the change is significant in absolute terms.

Change matrices of the original data provide a compact visualization of all the observed changes. The decomposition components show additional information and allow the detection of small relative changes in each component.

2) *Wheat*: Wheat [see Fig. 10(b)] was still growing at the beginning of the campaign and reached its full height shortly after the first acquisitions. The dashed line indicates the harvest.



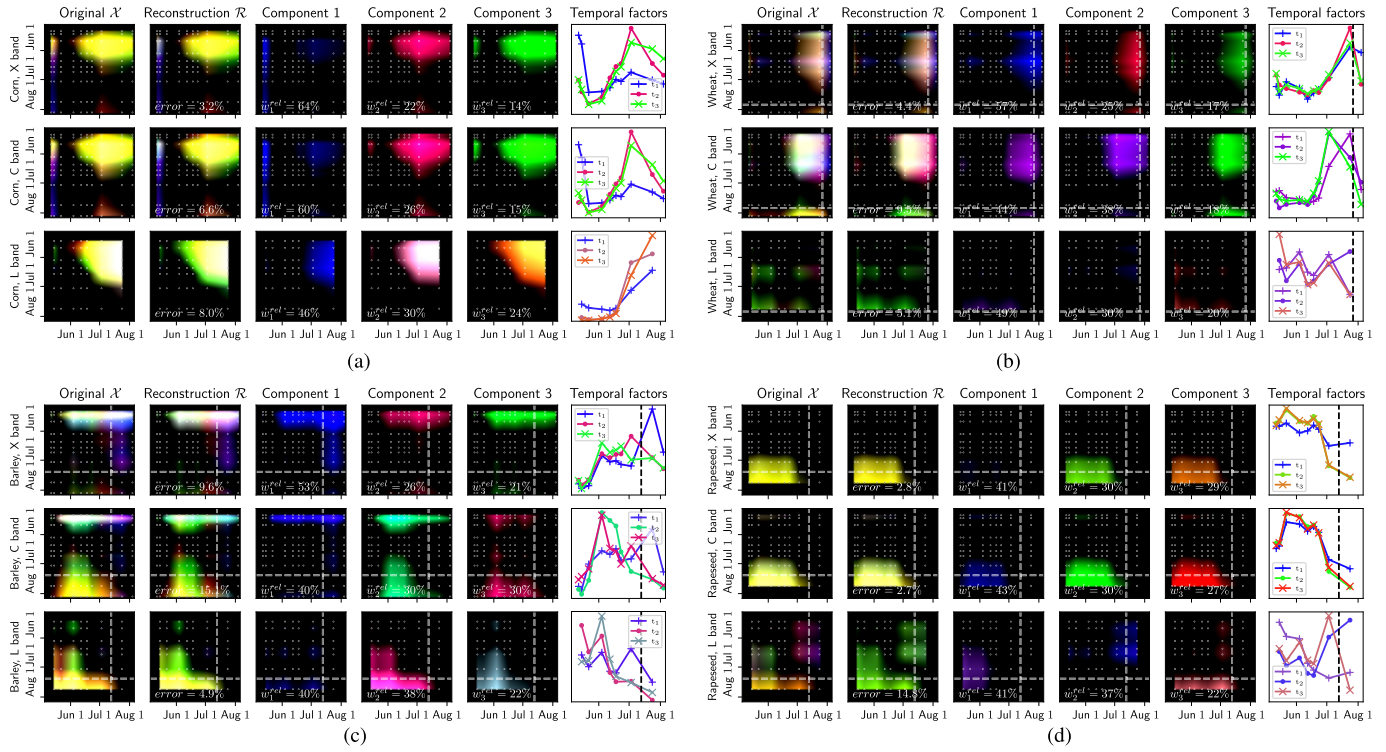


Fig. 10. Decomposition results for different crops at the X-, C-, and L-bands. Dashed lines indicate crop harvest. (a) Corn. (b) Wheat. (c) Barley. (d) Rapeseed.

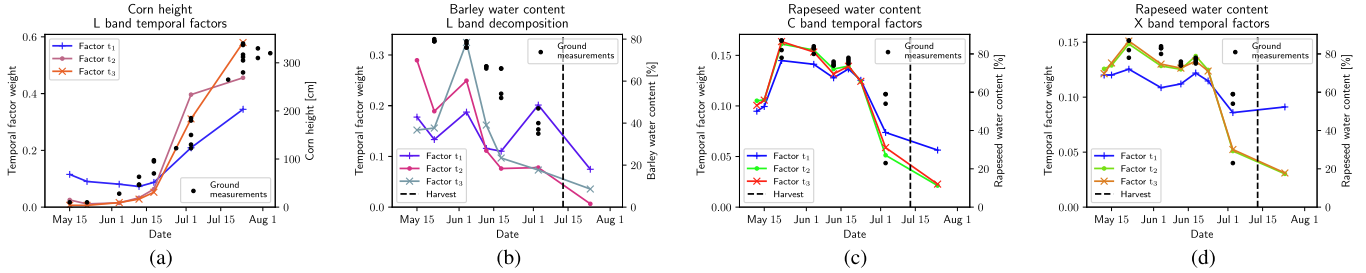


Fig. 11. Measured parameters of several crops show similar trends to the temporal factors. For the barley field, the measurements were obtained from a similar field nearby for the first two dates. Dashed lines indicate crop harvest. (a) Corn at the L-band and the measured crop height. (b) Barley at the L-band and measured plant water content. (c) Rapeseed at the C-band and measured plant water content. (d) Rapeseed at the X-band and measured plant water content.

The X-band shows no significant changes till the end of the campaign. All the components show a strong increase just before the harvest and a decrease after. The decrease is less pronounced in the first component related to the surface.

The C-band also shows a strong increase; however, it happens earlier during the campaign around the fruit maturation period. The first two polarimetric factors of the C-band show a mixture of surface and double-bounce scattering.

The L-band signal increases and decreases over time with no clear structure. In contrast to the X- and C-bands, the changes caused by fruit maturation and plant drying are not so significant for the L-band. The changes mostly occur at the head of the plant which is mostly transparent to the L-band as the head is smaller than the wavelength.

For wheat, no clear similarities between the temporal factors and the ground measurements were observed. A possible reason is that several processes (such as drying and geometrical changes) counteract each other and cannot be separated by the decomposition. In this case, extending the decomposition by

another data dimension (e.g., interferometry) might improve the separation.

3) *Barley*: Barley [see Fig. 10(c)] has reached its maximal height at the beginning of the campaign and the observed changes are related to fruit maturation, drying, and harvest.

The X-band starts with a low signal power and increases during fruit maturation. The drying and the bending of the head and the stalks in June do not show large changes. The harvest is visible as a strong increase in the surface component (blue) and a small decrease in dihedral (red) and volume (green) components.

The C-band also starts with a weak signal and shows a strong increase in the second and the third components in the middle of the acquisition period. Toward the later dates, drying and harvest are responsible for the strong decrease visible in the lower triangular part of the change matrices. Interestingly, the second component appears in green indicating a slightly stronger volume scattering than dihedral effects represented by the third component.

The L-band shows a decrease in all the components except the first (surface scattering) which shows alternating increase and decrease over time. The decrease in the second and third components can be attributed to the drying process, as the crops become increasingly transparent to the long-wavelength radar signal. The decreasing dynamics are similar to the measured volumetric water content as shown in Fig. 11(b). Note that the first two ground measurements were taken from a different field and do not necessarily match the conditions on the decomposed field very well.

4) *Rapeseed*: Similar to barley, the rapeseed [see Fig. 10(d)] plants were already fully grown before the first acquisition.

The X- and C-bands show similar changes over time. The backscatter remains strong in the first half of the campaign until the plants start to dry out. The changes are best seen in the second and third components which are mostly related to volume and dihedral scattering. The strongest component related to surface scattering shows very little change at the X-band with an almost completely black change matrix.

The temporal factors at the C- and X-bands show a correlation to the plant water content in Fig. 11(c) and (d). Both the second and third factor shapes are similar to the measured plant water content. The separation between the first factor and the others is less pronounced at the C-band than at the X-band.

The L-band signal shows complex dynamics with several changes. The similarities to the change matrices of wheat hint that similar processes including drying are observed in both the cases. The first component steadily decreases over time, the second component shows an increase closer to the drying and the harvest phases, and the third component has several increases and decreases.

#### D. Comparison of Different Fields

To evaluate the stability of the proposed decomposition, we compare the results obtained from different fields. While each field can have different conditions (irrigation, soil type, etc.), we expect the overall dynamics to be similar.

Fig. 12 shows the decomposition output for five different corn fields observed at the L-band. The reconstruction error is around 8%–9% for all the fields. The first component is dominated by the surface scattering and has a similar polarimetric signature across all the fields. The second and third components show a mix of scattering mechanisms with dihedral scattering being the strongest. While polarimetry shows differences, the temporal factors of the second and third components have similar trends across different fields.

Corn field 5 shows differences in the shape of the first temporal factor compared with the other fields. The polarimetric signature indicates a clear surface scattering contribution that is interpreted as a change in the soil conditions. To confirm this hypothesis, we have investigated the available polarimetric images for fields 4 and 5. Both the fields are located next to each other and have a minimal difference in the incidence angle. Field 4 shows a change in the soil structure between the first two acquisitions. The row structures are clearly visible during the first acquisition, but they are not pronounced during the second acquisition resulting in less backscatter.

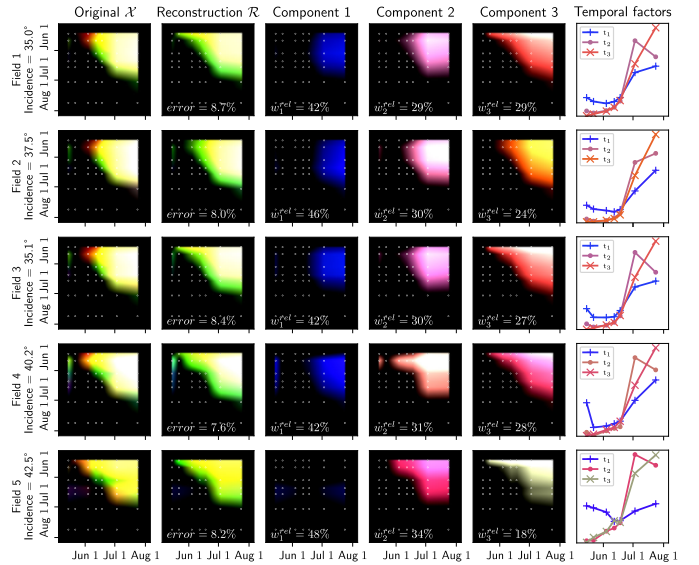


Fig. 12. Comparison of different corn fields at the L-band.

Field 5 shows no significant changes: row structures are visible on both the acquisitions. This different behavior is directly visible in the corresponding change matrix of the first component of field 5 with no significant increase or decrease. In contrast, the temporal factors of the second and third components show similar behavior across all the fields. This example illustrates that the proposed technique can isolate different temporal trends into different factors.

## IV. DISCUSSION

Polarimetric time series decomposition is a useful tool for crop analysis that separates the polarimetric signal into components and allows to analyze the changes for each component individually. In this section, we highlight the potential benefits of decomposition for crop parameter estimation, discuss connections to existing polarimetric decompositions, and propose possible improvements to the general constrained decomposition framework.

#### A. Toward Crop Parameter Estimation

Crop and cropland parameter estimation using SAR data is an active research area. Recently, there has been a lot of progress in land cover and crop classification using supervised machine learning methods [38], [39], [40]. However, the estimation of parameters such as plant height or water content with supervised machine learning approaches is still challenging due to the very limited amount of available ground measurements. Other approaches for crop parameters' estimation, especially for crop height, often use techniques adapted from forest height estimation [41], [42]. In contrast to forests, crop height estimation is more challenging as crops tend to change faster. Higher precision in the estimation is important and the existing techniques still require improvements for practical applicability [43].

1) *Crop-Specific Polarimetric Models*: Polarimetric models have been proposed to estimate physical parameters including soil moisture over vegetated agricultural areas [44]. Typically,

the same model is inverted regardless of the crop type. Since different crops show large differences in the geometrical structure, growth dynamics, and maximal height, a single polarimetric model is unlikely to perform well across all the crop types. Different crop-specific polarimetric models are likely to estimate the parameters more accurately.

Tensor decompositions reduce the data dimensionality and help with exploratory data analysis for individual crop types. The polarimetric time series decomposition is not limited to a single field but can be used to extract the main polarimetric mechanisms across many different fields. The coherency matrices from all the acquisition dates and fields with the same crop type are stacked along the temporal axis and jointly decomposed. The resulting polarimetric factors describe the main scattering mechanisms and can help develop crop-specific physical models. For example, an appropriate volume scattering representation for each crop type can be selected based on the polarimetric factors. Combined with crop classification, an appropriate polarimetric model can then be inverted for each field.

2) *Integration of Prior Knowledge:* The estimation accuracy of crop parameters can be improved by integrating prior knowledge as shown in [45], where a logistic growth model assists the height estimation of rice. In the proposed decomposition framework, prior knowledge can be integrated by adding additional constraints to the temporal factors. For example, the factors can be constrained to follow the expected parameter dynamics, such as increasing biomass or decreasing plant water content. The expected dynamics can be extracted from crop models such as decision support system for agrotechnology transfer (DSSAT) [46] that simulate the evolution of crop and soil parameters in time. The most promising temporal factors to be constrained are those already showing a correlation to the crop parameters as shown in Fig. 11.

### B. Connection to Polarimetric Decompositions

The proposed tensor decomposition for polarimetric time series has a few notable differences from the classical polarimetric decompositions (e.g., Cloude and Pottier [2], Freeman and Durden [3], or Yamaguchi et al. [5]). The main difference is the input: while classical decompositions operate on a single coherency matrix, the proposed decomposition operates on a tensor created from multiple coherency matrices. Furthermore, the outputs differ significantly. Classical decompositions produce a set of parameters that characterize the input coherency matrix. In contrast, the proposed decomposition produces a set of factors where each of the factors corresponds to a data dimension. This allows the analysis of the factors with techniques specific to the dimension and simplifies data understanding. For example, one of the dimensions in the proposed decomposition is polarimetry with the corresponding polarimetric coherency matrices as factors. In that sense, classical polarimetric decompositions can be applied to the results of the proposed decomposition.

In this article, we refer to the canonical Pauli scattering mechanisms to interpret the polarimetric factors. When each polarimetric factor describes exactly one individual Pauli mechanism (component change matrices appearing in pure red,

green, and blue), the temporal factor trends are very close to the trends of the Pauli elements in the original tensor. This is more common at shorter wavelengths, especially at the X-band. At longer wavelengths (L-band), the polarimetric factors often represent a mix of Pauli polarizations and the temporal factors show different dynamics from the time series of Pauli elements in the original tensor. More research is needed to explain these observations.

### C. Possible Extensions

Tensor decompositions on multidimensional SAR data is a new research area and opens several directions for exploration. We discuss possible extensions to the constrained decomposition framework and ways to improve the existing decompositions or integrate additional data dimensions using the framework.

1) *Constrained SKP Decomposition:* The SKP decomposition introduced by Tebaldini [7] extracts polarimetric and structural factors from a multibaseline multipolarization matrix. The original implementation uses SVD in combination with several reshaping operations and component recombinations to obtain the final factors. The SVD performs the decomposition on vectorized matrices and does not necessarily preserve some matrix properties. Therefore, component recombination is required to ensure that the final factors are PSD and carry a physical meaning. The component recombination parameters typically have a range of valid values meaning that the recombination solution is not unique. In cases where the data do not follow the assumptions from [7], the recombination can fail yielding no solution. Furthermore, the original implementation only allows to obtain two components.

We can increase the robustness of the decomposition, allow an arbitrary number of components, and use different loss functions using the constrained decomposition framework. First, we constrain the factors to be PSD. Then, we reconstruct the approximation with the sum of Kronecker products. Finally, we compute the loss and optimize until convergence. The factors are guaranteed to be PSD by construction. In addition, we can address solution uniqueness by adding stricter constraints and limiting the rank of the factor matrices.

2) *Integration of Physical Models:* Physical models have been used for the interpretation of SAR data and parameter inversion [47], [48]. The constrained decomposition framework can be integrated with the models for one or several data dimensions. The key observation is that any differentiable function can be used for reconstruction or factor constraint implementation. Therefore, we can integrate any physical model that is represented as a differentiable function reconstructing a factor from a set of physical parameters. For example, a polarimetric matrix factor associated with a specific physical process (e.g., scattering from a surface or a volume of particles) can be reconstructed from the parameters defining the process. After the optimization procedure, the decomposition results provide the physical parameters that best describe the data.

The physical model can also be seen as a specific type of constraint that limits the solution space to the factor subset spanned by the model. Physical models can be applied to

individual data dimensions and combined with the standard factors describing data dimensions with no known models.

3) *Computational Performance*: The performance of the decomposition framework is important for practical applications. Depending on the size and dimensionality of the tensor, the optimization procedure can take many steps until convergence. In our implementation, we rely on the PyTorch library to run the decomposition efficiently. PyTorch supports optimization on GPUs further speeding up heavy matrix operations.

The decomposition performance can further be improved using a suitable optimizer. When the size of the factors and the number of free parameters are relatively small, second-order optimizers such as BFGS typically lead to fewer optimization steps and faster convergence. Several optimizers are already implemented in PyTorch and can be used for decomposition optimization.

Finally, the initialization of the factors has a large effect on the convergence speed. Factors initialized to values using prior knowledge about the solution require fewer optimization steps until convergence.

## V. CONCLUSION

This work introduces a constrained tensor decomposition framework for SAR data. We generalize the CP decomposition, extend it with constraints, and allow different factor shapes and loss functions. The framework enables joint analysis of multidimensional SAR data, can be integrated into existing SAR decompositions, and opens ways to connect data-driven and model-based approaches by integrating physical models. We formulate the decomposition as an optimization problem and implement it using PyTorch, a well-established machine learning and optimization framework. Automatic differentiation and optimization simplify the decomposition design, facilitate experiments with different data dimensions, and allow to concentrate on the choice of the constraints and interpretation of the factors. The importance of constraints for physical validity, interpretability, and solution uniqueness has been discussed.

To demonstrate the decomposition framework, we formulate the polarimetric time series decomposition, apply it to agricultural data, and analyze the development of four different crop types. The decomposition reduces the dimensionality and extracts the polarimetric and temporal factors that best explain the data. Notable differences to the existing SAR decompositions include nonorthogonal factors and a variable number of components.

We extend the polarimetric change analysis visualization with the decomposition results to show additional details. By performing the decomposition and plotting a change matrix for each component, we can detect more small and fine-grained changes in each component. This approach also ensures that large relative changes in weak components do not hide small changes in strong components. The component weights provided by the decomposition indicate the relative importance of each component.

The polarimetric time series decomposition is not limited to agricultural areas and can be used for many other applications.

The decomposition describes the time series in a compact way: polarimetric factors define the scattering mechanism for each component while temporal factors describe how the intensity of that mechanism changes over time. The information can be used directly for exploratory data analysis or be combined with other methods including change analysis, as presented in this article.

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for giving valuable suggestions. Special thanks to Prof. Dr. Senthold Asseng for academic advice and discussions.

## REFERENCES

- [1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 1, pp. 6–43, Mar. 2013.
- [2] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 68–78, Jan. 1997.
- [3] A. Freeman and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 963–973, May 1998.
- [4] E. Krogager, "New decomposition of the radar target scattering matrix," *Electron. Lett.*, vol. 26, no. 18, p. 1525, 1990.
- [5] Y. Yamaguchi, T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric SAR image decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1699–1706, Aug. 2005.
- [6] S.-W. Chen, Y.-Z. Li, X.-S. Wang, S.-P. Xiao, and M. Sato, "Modeling and interpretation of scattering mechanisms in polarimetric synthetic aperture radar: Advances and perspectives," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 79–89, Jul. 2014.
- [7] S. Tebaldini, "Algebraic synthesis of forest scenarios from multibaseline PolInSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4132–4142, Dec. 2009.
- [8] H. Abdi, "The eigen-decomposition: Eigenvalues and eigenvectors," in *Encyclopedia of Measurement and Statistics*, N. J. Salkind, Ed. Newbury Park, CA, USA: Sage, 2007, pp. 304–308. [Online]. Available: <https://us.sagepub.com/en-us/nam/encyclopedia-of-measurement-and-statistics/book227214> and <https://personal.utdallas.edu/~herve/Abdi-EVD2007-pretty.pdf>
- [9] C. D. Martin and M. A. Porter, "The extraordinary SVD," *Amer. Math. Monthly*, vol. 119, no. 10, p. 838, 2012.
- [10] I. T. Jolliffe, "Principal component analysis for special types of data," in *Principal Component Analysis*. Cham, Switzerland: Springer, 2002.
- [11] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 13, 2000, pp. 1–7.
- [12] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of 'Eckart–Youn' decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, Sep. 1970.
- [13] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multi-model factor analysis," UCLA Work. Papers Phonetics, 16, UCLA Phonetics Lab, UCLA College, Univ. California, Los Angeles, 1970. [Online]. Available: <http://phonetics.linguistics.ucla.edu/workpapph/wpp.htm> and <https://escholarship.org/uc/item/0410x385>
- [14] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1966.
- [15] P. M. Kroonenberg, *Applied Multiway Data Analysis*. Hoboken, NJ, USA: Wiley, 2008.
- [16] A. Shashua and A. Levin, "Linear image coding for regression and classification using the tensor-rank principle," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Dec. 2001, p. 1.
- [17] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
- [18] M. Wang et al., "Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 1, pp. 26–72, Mar. 2023.

- [19] M. Tao, F. Zhou, Y. Liu, and Z. Zhang, "Tensorial independent component analysis-based feature extraction for polarimetric SAR data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2481–2495, May 2015.
- [20] J. Kang, Y. Wang, and X. X. Zhu, "Multipass SAR interferometry based on total variation regularized robust low rank tensor decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5354–5366, Aug. 2020.
- [21] J. Luo et al., "Despeckling multitemporal polarimetric SAR data based on tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 9858–9873, 2023.
- [22] A. Alonso-González, C. López-Martínez, K. P. Papathanassiou, and I. Hajnsek, "Polarimetric SAR time series change analysis over agricultural areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 10, pp. 7317–7330, Oct. 2020.
- [23] S. R. Cloude, "A physical approach to PolSAR time series change analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–4, 2022.
- [24] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 498–518, Mar. 1996.
- [25] J. S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL, USA: CRC Press, 2017.
- [26] D. Hong, T. G. Kolda, and J. A. Dersch, "Generalized canonical polyadic tensor decomposition," *SIAM Rev.*, vol. 62, no. 1, pp. 133–163, Jan. 2020.
- [27] M. Lezcano Casado, "Trivializations for gradient-based optimization on manifolds," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–12.
- [28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [29] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.*, vol. 16, no. 5, pp. 1190–1208, Sep. 1995.
- [30] R. Fletcher, *Practical Methods of Optimization*. Hoboken, NJ, USA: Wiley, 2013.
- [31] A. Paszke et al., "Automatic differentiation in PyTorch," in *Proc. 31st Conf. Neural Inf. Process. Syst.* 2017, pp. 1–4.
- [32] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Design Implement. (OSDI)*, Nov. 2016, pp. 265–283.
- [33] J. Bradbury et al. *JAX: Composable Transformations of Python+NumPy Programs*. Accessed: Oct. 13, 2023. [Online]. Available: <http://github.com/google/jax>
- [34] A. Gulli and S. Pal, *Deep Learning With Keras*. Birmingham, U.K.: Packt Publishing, 2017.
- [35] A. Marino, S. R. Cloude, and J. M. Lopez-Sanchez, "A new polarimetric change detector in radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2986–3000, May 2013.
- [36] R. Horn, A. Nottensteiner, A. Reigber, J. Fischer, and R. Scheiber, "F-SAR—DLR's new multifrequency polarimetric airborne SAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 2, Jul. 2009, pp. 902–905.
- [37] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
- [38] H. Parikh, S. Patel, and V. Patel, "Classification of SAR and PolSAR images using deep learning: A review," *Int. J. Image Data Fusion*, vol. 11, no. 1, pp. 1–32, Jan. 2020.
- [39] Z. Sun, D. Wang, and G. Zhong, "A review of crop classification using satellite-based polarimetric SAR imagery," in *Proc. 7th Int. Conf. Agro-Geoinformatics (Agro-geoinformatics)*, Aug. 2018, pp. 1–5.
- [40] M. Arias, M. Á. Campo-Bescós, and J. Álvarez-Mozos, "Crop classification based on temporal signatures of Sentinel-1 observations over Navarre province, Spain," *Remote Sens.*, vol. 12, no. 2, p. 278, Jan. 2020.
- [41] C.-A. Liu, Z.-X. Chen, Y. Shao, J.-S. Chen, T. Hasi, and H.-Z. Pan, "Research advances of SAR remote sensing for agriculture applications: A review," *J. Integrative Agricult.*, vol. 18, no. 3, pp. 506–525, Mar. 2019.
- [42] E. Erten, J. M. Lopez-Sanchez, O. Yuzugullu, and I. Hajnsek, "Retrieval of agricultural crop height from space: A comparison of SAR techniques," *Remote Sens. Environ.*, vol. 187, pp. 130–144, Dec. 2016.
- [43] N. Romero-Puig and J. M. Lopez-Sanchez, "A review of crop height retrieval using InSAR strategies: Techniques and challenges," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7911–7930, 2021.
- [44] I. Hajnsek, T. Jagdhuber, H. Schon, and K. P. Papathanassiou, "Potential of estimating soil moisture under vegetation cover by means of PolSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 442–454, Feb. 2009.
- [45] N. Li et al., "Rice crop height inversion from TanDEM-X PolInSAR data using the RVoG model combined with the logistic growth equation," *Remote Sens.*, vol. 14, no. 20, p. 5109, Oct. 2022.
- [46] J. W. Jones et al., "The DSSAT cropping system model," *Eur. J. Agronomy*, vol. 18, nos. 3–4, pp. 235–265, Jan. 2003.
- [47] I. Hajnsek, E. Pottier, and S. R. Cloude, "Inversion of surface parameters from polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 4, pp. 727–744, Apr. 2003.
- [48] S. R. Cloude and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *IEE Proc. Radar, Sonar Navigat.*, vol. 150, no. 3, p. 125, Jun. 2003.



**Nikita Basargin** received the B.Sc. and M.Sc. degrees in computer science from the Technical University of Munich (TUM), Munich, Germany, in 2017 and 2020, respectively. He is currently pursuing the Ph.D. degree in cooperation with the Munich School for Data Science (MUDS) and the TUM School of Life Sciences, TUM, Freising, Germany.

Since 2021, he has been with the Microwaves and Radar Institute, German Aerospace Center (DLR), Wessling, Germany. His research interests include application of data science and machine learning methods to remote sensing SAR data, visualization, optimization, and biophysical information retrieval.

Mr. Basargin was a member of the Max Weber Program from 2014 to 2020. He was a recipient of the Rohde & Schwarz Best Bachelor Award in 2017 and the IEEE GRSS IADF School Best Poster Award in 2023.



**Alberto Alonso-González** received the B.Sc. degree in computer science, the M.Sc. degree in telecommunication engineering, and the Ph.D. degree from the Technical University of Catalonia (UPC), Barcelona, Spain, in 2007, 2009, and 2014, respectively.

From 2009 to 2014, he was with the Department of Signal Theory and Communications, UPC. From 2014 to 2022, he was with the Microwaves and Radar Institute, German Aerospace Center (DLR), Wessling, Germany, as a member of the Polarimetric SAR Interferometry Research Group. Since 2022, he has been with the CommSensLab Group, Department of Signal Theory and Communications, UPC. His research interests include multidimensional synthetic aperture radar (SAR), SAR polarimetry and interferometry, and digital signal and image processing. His research activities are focused on remote sensing data modeling and machine learning techniques applied to agricultural and forest monitoring for climate change mitigation.

Dr. Alonso-González was a recipient of the First Place Student Paper Award at the EUSAR 2012 Conference and the Extraordinary Doctoral Thesis Award from the Technical University of Catalonia in 2016.



**Irena Hajnsek** (Fellow, IEEE) received the Diploma degree (Hons.) in fluvial river systems from the Free University of Berlin, Berlin, Germany, in 1996, and the Dr.rer.nat. degree (Hons.) in model-based estimation of soil moisture from fully polarimetric synthetic aperture radar from the Friedrich Schiller University of Jena, Jena, Germany, in 2001.

Since November 2009, she has been a Professor of Earth observation at the Institute of Environmental Engineering, Swiss Federal Institute of Technology (ETH) Zürich, Zürich, Switzerland, and at the same time, the Head of the Polarimetric SAR Interferometry Research Group, Microwaves and Radar Institute, German Aerospace Center (DLR), Wessling, Germany. Since 2010, she has been the Science Coordinator of the German satellite mission TanDEM-X. She is currently a member of the European Space Agency Mission Advisory Group for the ROSE-L Mission. Her main research interests include electromagnetic propagation and scattering theory, radar polarimetry, SAR and interferometric SAR data processing techniques, and environmental parameter modeling and estimation.

Dr. Hajnsek was an Elected Member of the IEEE GRSS AdCom from 2013 to 2021. She was the Technical Program Co-Chair for the IEEE IGARSS 2012, Munich, Germany, and the IEEE IGARSS 2019, Yokohama, Japan. She was the Vice President of the IEEE GRSS Technical Committees from 2016 to 2020.