HyperLISTA-ABT: An Ultralight Unfolded Network for Accurate Multicomponent Differential Tomographic SAR Inversion

Kun Qian, Yuanyuan Wang, *Member, IEEE*, Peter Jung, *Member, IEEE*, Yilei Shi, *Member, IEEE*, and Xiao Xiang Zhu, *Fellow, IEEE*

***Abstract*— Deep neural networks based on unrolled iterative algorithms have achieved remarkable success in sparse reconstruction applications, such as synthetic aperture radar (SAR) tomographic inversion (TomoSAR). However, the currently available deep learning-based TomoSAR algorithms are limited to 3-D reconstruction. The extension of deep learning-based algorithms to 4-D imaging, i.e., differential TomoSAR (D-TomoSAR) applications, is impeded mainly due to the high-dimensional weight matrices required by the network designed for D-TomoSAR inversion, which typically contain millions of freely trainable parameters. Learning such huge number of weights requires an enormous number of training samples, resulting in a large memory burden and excessive time consumption. To tackle this issue, we propose an efficient and accurate algorithm called HyperLISTA-ABT. The weights in HyperLISTA-ABT are determined in an analytical way according to a minimum coherence criterion, trimming the model down to an ultra-light one with only three hyperparameters. Additionally, HyperLISTA-ABT improves the global thresholding by utilizing an adaptive block-wise thresholding (ABT) scheme, which applies block-coordinate techniques and conducts thresholding in local blocks, so that weak expressions and local features can be retained in the shrinkage step layer by layer. Simulations were performed and demonstrated the effectiveness of our approach, showing that HyperLISTA-ABT achieves superior computational efficiency with no significant performance degradation compared to the state-of-the-art methods. Real data experiments showed that a high-quality 4-D point cloud could be reconstructed over a large area by the proposed HyperLISTA-ABT with affordable computational resources and in a fast time.**

Manuscript received 9 July 2023; revised 7 December 2023 and 26 March 2024; accepted 10 April 2024. Date of publication 25 April 2024; date of current version 2 May 2024. This work was supported in part by the European Union (ERC-PoC), [Artificial Intelligence for Smart Cities (AI4SmartCities)], under Grant 957467; in part by the German Federal Ministry of Education and Research in the framework of the international future AI Laboratory "AI4EO–Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond" under Grant 01DD20001; in part by the Excellence Strategy of the Federal Government and the Länder through the TUM Innovation Network EarthCare; and in part by the Munich Center for Machine Learning. *(Corresponding author: Xiao Xiang Zhu.)*

Kun Qian and Yuanyuan Wang are with the Chair Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany (e-mail: kun9361.qian@tum.de; y.wang@tum.de).

Peter Jung was with the Chair Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany (e-mail: peter.jung@tum.de).

Yilei Shi is with the School of Engineering and Design, Technical University of Munich, 80333 Munich, Germany (e-mail: yilei.shi@tum.de).

Xiao Xiang Zhu is with the Chair Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany, and also with the Munich Center for Machine Learning, 80333 Munich, Germany (e-mail: xiaoxiang.zhu@tum.de).

Digital Object Identifier 10.1109/TGRS.2024.3391066

***Index Terms*— Differential synthetic aperture radar (SAR) tomography (D-TomoSAR), HyperLISTA, sparse recovery, unrolling algorithms.**

NOMENCLATURE

$\boldsymbol{\gamma}^k$	Output at the kth layer.
$\mathbf{R}$	Original steering matrix.
$\mathbf{W}$	Optimized weight matrix.
$\mathbf{g}$	Complex-valued SAR acquisitions.
$\eta_{\boldsymbol{\theta}^k}^{p^k}$	Soft-thresholding function with support selection at the kth layer.
p^k	Percentage for support selection at the kth layer.
$\boldsymbol{\theta}^k$	Threshold controlling the soft-thresholding function at the kth layer.
β^k	Factor balancing gradient and momentum.
c_1, c_2, c_3	Hyperparameters in HyperLISTA.
h_1, h_2, h_3	Hyperparameters in HyperLISTA-ABT.

I. INTRODUCTION

SYNTHETIC aperture radar tomography (TomoSAR) has attracted significant interest due to its capability in 3-D reconstruction, particularly for urban areas [1], [2], [3], [4]. Compressive sensing [5], [6] (CS)-based algorithms are usually preferred for solving TomoSAR inversion [7], [8], [9]. However, the heavy computational cost of CS-based methods makes them less applicable for large-scale processing. Among the different methods aiming to tackle this issue, deep neural networks have been employed in speeding up TomoSAR inversion. In the work presented in [10], TomoSAR inversion was approached as a classification problem, and a conventional convolutional neural network (CNN) was employed to solve the problem. However, this approach was limited to the detection of single scatterers, and it did not fully address the challenges of TomoSAR inversion for complex scenes with multiple scatterers and variations in the elevation direction. More recently, thanks to an emerging deep learning technique called deep unfolding [11], the authors proposed $\boldsymbol{\gamma}$-Net in [12] for improving the unrolled iterative shrinkage thresholding algorithm (ISTA) network. It was shown that $\boldsymbol{\gamma}$-Net could succeed in accelerating the processing speed by 2–3 order of

magnitude while maintaining a comparable super-resolution power and location accuracy compared to second-order CS solvers. In addition, a gated recurrent structure, dubbed as complex-valued sparse minimal gated units (CV-SMGUs), was proposed in [13] that incorporates historical information into the dynamics of network, thus preserving the full information. As discussed in [13], CV-SMGUs could outperform γ-Net by a fair margin.

However, to the best of our knowledge, deep learning-based TomoSAR algorithms are to date still confined to 3-D reconstruction cases. Considering that spaceborne datasets are usually acquired in the repeat-pass mode at different time stamps, often over several years, it is necessary to additionally account for a potential deformation of objects in the estimation, such as seasonal motion caused by thermal dilation or linear motions like subsidence. The 4-D imaging technique taking into account additional deformation parameters is known as D-TomoSAR [1], [14], [15], [16].

The limitation of deep learning-based algorithms in solving D-TomoSAR inversion is mainly attributed to the high-dimensional weight matrices to be learned in the network. For modern deep learning-based algorithms, like γ-Net and CV-SMGUs, the size of the weight matrices is usually related to the discretization level. In D-TomoSAR cases, especially when multicomponent motion terms are considered, weight matrices can easily contain over one million free trainable parameters. As a consequence, it would be extremely computationally inefficient to learn those weights without mentioning the enormous number of training samples required. A detailed analysis of this issue can be found in Section III-A of the present article.

To tackle the computational challenges posed by learning huge weights, a pioneering solution was introduced in the seminal work by Liu et al. [17]. In their research, they proposed an analytical weight determination method. This method was further refined by them and extended into HyperLISTA [18]. By employing analytical weight determination, a novel perspective emerges for leveraging deep learning-based algorithms in D-TomoSAR inversion. Specifically, the optimization of weights using a data-free approach becomes possible, circumventing the need for an extensive number of training samples. This innovative approach effectively overcomes the limitations that were previously encountered in deep learning-based D-TomoSAR inversion.

However, it is important to note that, similar to LISTA, HyperLISTA employs a global thresholding scheme where a unified threshold is used to prune all entries. The choice of an appropriate threshold is thus crucial. A high threshold value may result in the loss of significant information [13], such as local features generated by echoed signals from dark scatterers. On the other hand, a low threshold value can delay convergence and yield a solution that lacks sufficient sparsity.

A. Contribution of the Present Study to the Field

To overcome the aforementioned issue in D-TomoSAR inversion, we proposed an ultralight model, named HyperLISTA-ABT, that improves HyperLISTA [18] through incorporating an ABT scheme. Same as HyperLISTA, the proposed HyperLISTA-ABT can be viewed as an unrolled ISTA network, whereas the weight matrices therein can be determined with analytical optimization according to the minimum coherence criterion. A system matrix with low mutual coherence implies a recovery of high probability, which is the fundamental concept of CS. The ABT scheme in HyperLISTA-ABT enables updating the block coordinates and conducting a shrinkage in local regions. Moreover, the blocksize is adjusted layer by layer for a better fine-focusing ability. The main contribution of this article is listed as follows.

1) We propose the efficient and accurate algorithm HyperLISTA-ABT and, to the best of our knowledge, are the first to apply deep neural networks to solve D-TomoSAR and multicomponent D-TomoSAR inversion.
2) We apply a block-coordinate technique and propose an ABT scheme to replace global thresholding in most shrinkage thresholding methods. Therefore, the local features from a weakly echoed signal can be possibly retained.
3) We carry out a systematic performance evaluation using both simulated and real data. The results demonstrate that the proposed HyperLISTA-ABT provides competitive estimation accuracy and superior computational efficiency. Large-scale D-TomoSAR processing was conducted, demonstrated by a 4-D point cloud reconstruction over Las Vegas.

The remainder of the article is outlined as follows. The high-dimensional SAR imaging model for D-TomoSAR and the limitation of deep learning-based algorithms in solving D-TomoSAR inversion are discussed in Section II. Section III provides an overview of the formulation of the proposed HyperLISTA-ABT with application to TomoSAR inversion. Results of systematic evaluation using simulated data and the practical demonstration are presented in Sections IV and V. In addition, further discussion on key differencing features of the proposed HyperLISTA-ABT is provided in Section VI. Finally, the conclusion of this article is drawn in Section VI. A brief overview of key symbols used throughout the article is listed in the Nomenclature.

II. Background

A. High-Dimensional SAR Imaging Model for D-TomoSAR

D-TomoSAR employs multibaseline and multitemporal synthetic aperture radar (SAR) acquisitions to estimate scattering profiles. Based on the estimated scattering profiles, we can reconstruct the 3-D distribution of scatterers along the elevation direction and the motion history assigned to each elevation position [1], [14], [15]. The imaging geometry can be seen in Fig. 1. The following describes the D-TomoSAR imaging model:

$$g_n = \int_{\Delta s} \boldsymbol{\gamma}(s) \exp(-j2\pi(\xi_n s + 2d(s, t_n)/\lambda))\mathrm{d}s \tag{1}$$

where g_n is the complex-valued SAR acquisition at a certain azimuth-range pixel at time t_n $(n = 1, 2, \ldots, N)$; $\boldsymbol{\gamma}(s)$

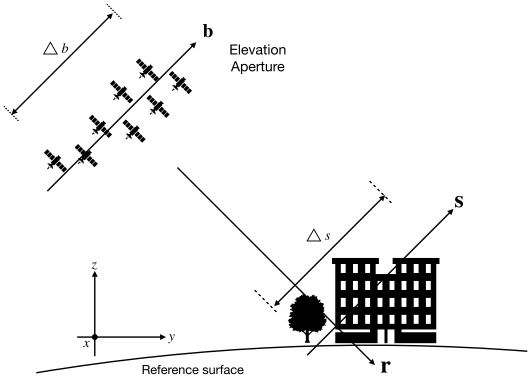


Fig. 1. SAR imaging geometry at a fixed azimuth position. The elevation synthetic aperture is built up by acquisition from slightly different incidence angles. Flight direction is orthogonal into the plane.

denotes the scattering profile along the elevation direction with an extent of $\triangle s$; $\xi = 2b_n/\lambda r$ is the elevation frequency proportional to the respective aperture position b_n; and $d(s, t_n)$ depicts the line-of-sight (LOS) motion, which is a function of elevation and time. The LOS motion relative to the master acquisition can be modeled with a linear combination of M base functions $\tau_M(t_n)$

$$d(s, t_n) = \sum_{m=1}^{M} p_m(s)\tau_m(t_n) \tag{2}$$

where $p_m(s)$ is the corresponding motion coefficient to be estimated. The choice of the base functions $\tau_m(t_n)$ depends on the underlying physical motion processes. Great details about how to choose proper base functions can be found in [16]. Taking multicomponent motion into consideration, we generalize (1) as follows:

$$\begin{aligned} g_n = \int \cdots \int \int \gamma(s)\delta(p_1 - p_1(s), \ldots, p_M - p_M(s)) \\ \times \exp\left(j2\pi\left(\xi_n s + \eta_{1,n} p_1 + \cdots + \eta_{M,n} p_M\right)\right) \\ \times d\ sdp_1, \ldots, d\ p_M. \end{aligned} \tag{3}$$

The inversion of the system model with multicomponent motion retrieves the elevation information as well as the motion history assigned to each elevation position, even if multiple scatterers are overlaid inside a resolution unit. Therefore, we can acquire a high-dimensional map of the scatterers. In the presence of noise $\boldsymbol{\varepsilon}$, the discrete high-dimensional D-TomoSAR model can be expressed as follows:

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{4}$$

where $\mathbf{g} \in \mathbb{C}^{N\times 1}$ is the complex-valued SAR measurement vector and $\mathbf{R} \in \mathbb{C}^{N\times L}$ is the irregular sampled Fourier transformation steering matrix, where N is the number of SAR acquisitions and L is the amount of the discretization in the signal to be reconstructed.

As investigated in [7], usually only a few (less than 4) scatterers are overlaid inside an individual pixel in urban areas, such that $\boldsymbol{\gamma}$ is sufficiently sparse so that retrieving $\boldsymbol{\gamma}$ can be formulated as a sparse reconstruction problem. Accordingly, solving $\boldsymbol{\gamma}$ in the presence of noise can be formulated as a

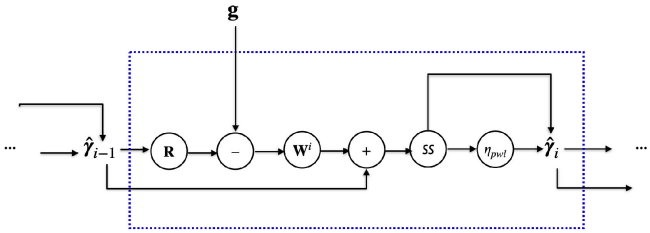


Fig. 2. Illustration of an intermediate layer in $\boldsymbol{\gamma}$-Net.

basis pursuit denoising (BPDN) optimization problem, which can be expressed as follows:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda\|\boldsymbol{\gamma}\|_1 \right\} \tag{5}$$

where λ is the regularization parameter controlling the data-fit terms and the signal sparsity. Great details about how to choose a proper value of λ according to the noise level can be found in [19].

B. Review of the Deep Learning-Based TomoSAR Algorithms and Their Limitation in Solving D-TomoSAR Inversion

In [12] and [13], the respective authors proposed two advanced deep learning-based algorithms by improving unrolled neural networks. Experimental results on both laboratory and real data demonstrated their strong super-resolution power and high location accuracy. However, their application to date is still limited to 3-D reconstruction. Taking $\boldsymbol{\gamma}$-Net as an example, we will explain the difficulty of applying deep learning-based algorithms for solving D-TomoSAR inversion. To start with, we briefly go through the basics of the $\boldsymbol{\gamma}$-Net architecture. Fig. 2 illustrates the structure of an intermediate layer of $\boldsymbol{\gamma}$-Net, which can be formally defined as follows:

$$\hat{\boldsymbol{\gamma}}^k = \eta_{\theta_k}^{\rho^k}\left\{\hat{\boldsymbol{\gamma}}^{k-1} + \mathbf{W}_k^H\left(\mathbf{g} - \mathbf{R}\hat{\boldsymbol{\gamma}}^{k-1}\right), \boldsymbol{\theta}_i\right\}. \tag{6}$$

More details about $\boldsymbol{\gamma}$-Net can be found in [12].

As we can see, in each $\boldsymbol{\gamma}$-Net layer, a weight matrix $\mathbf{W}_k$ of the size $N \times L$ needs to be learned. For 3-D reconstruction cases, the value of L is only determined by the grids number after the discretization of the elevation extent, thus it is typically in the range of hundreds for spaceborne sensors and $N \times L$ will be in thousands then. However, this number increases exponentially in D-TomoSAR cases when multicomponent motion terms, usually linear and periodic motions, are taken into consideration. The training of the network then becomes conversely a challenge due to the tremendous amount of free trainable parameters. For instance, when we consider two motion terms, i.e., linear and seasonal motion, the value of L will be determined by the product of the discretized grid numbers along each direction $L = L_s \times L_v \times L_a$, where L_s, L_v, and L_a indicate the discretization levels for elevation, linear motion, and seasonal motion, respectively. A very conservative level of discretization in elevation, linear motion, and periodic motion for TerraSAR-X image stacks L_s, L_v, and L_a would be 200, 50, and 50, respectively. When multiplied, the value of L will then be 0.5 million, meaning that, there will be millions of parameters to be learned in

each weight matrix. Such large weight matrices result in two unavoidable downsides. First, the model tends to converge at the ground truth instead of the LASSO minimizer, because the update direction $\mathbf{W}^H(\mathbf{g}-\mathbf{R}\gamma)$ does not align with the gradient of the l_2 term in the LASSO objective $\mathbf{R}^H(\mathbf{g}-\mathbf{R}\gamma)$. Therefore, we always need to train the model in a supervised way. Consequently, a massive number of training samples are required to train the model with huge weight matrices, thus making the training procedure extremely inefficient. Second, the training of the huge model requires a significant amount of GPU memory, which is usually not feasible with consumer-level GPUs.

III. Methodology

A. HyperLISTA With ABT (HyperLISTA-ABT)

To circumvent the tedious and troublesome model training caused by needing to learn huge weight matrices, an analytical weight optimization method, which is based on coherence minimization, was proposed in ALISTA [17] to determine the weights in an unrolled neural network designed for sparse recovery, such as LISTA. ALISTA combines the superior empirical performance of fully learned methods and significantly reduces the number of parameters, leaving only thresholds and stepsize parameters to be learned. In addition, an ultralight model, called HyperLISTA, was proposed in [18], which further trimmed down the training complexity. In HyperLISTA, weight matrices can be computed in a similar way to [17] and the training is reduced to tuning only three hyperparameters from the data. The following shows the formal update rules of HyperLISTA:

$$\boldsymbol{\gamma}^{k+1}=\eta_{\theta^k}^{p^k}(\boldsymbol{\gamma}^k+\boldsymbol{W}^H(\boldsymbol{g}-\boldsymbol{R}\boldsymbol{\gamma}^k)+\beta^k(\boldsymbol{\gamma}^k-\boldsymbol{\gamma}^{k-1})) \tag{7}$$

where

$$\theta^k=c_1\left\|\boldsymbol{R}^+\left(\boldsymbol{R}\boldsymbol{\gamma}^k-\boldsymbol{g}\right)\right\|_1 \tag{8}$$

$$\beta^k=c_2\left\|\boldsymbol{\gamma}_k\right\|_0 \tag{9}$$

$$p^k=c_3\min\left(\log\left(\frac{\left\|\boldsymbol{R}^+\boldsymbol{g}\right\|_1}{\left\|\boldsymbol{R}^+\left(\boldsymbol{R}\boldsymbol{\gamma}^k-\boldsymbol{g}\right)\right\|_1}\right),L\right) \tag{10}$$

where c_1, c_2, and c_3 indicate the three hyperparameters to be tuned. It is possible to learn the hyperparameters via backpropagation, albeit this method may be an overkill as it involves passing gradients through deep neural network layers to learn just three parameters. Less computationally expensive methods, such as grid search, could be employed to obtain a set of proper hyperparameters. Despite providing a bit less accurate estimate, the empirical findings in [18] showed that HyperLISTA is robust to perturbations in the values of $c1, c2$, and $c3$. In grid search, a coarse grid is first applied to find an interest region, and then this is zoomed-in with a fine-grained grid. The hyperparameters are determined by minimizing the normalized mean square error (NMSE) over the simulated ground truth. The NMSE is defined as follows:

$$\text{NMSE}=\frac{1}{T}\sum\frac{\|\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}\|_2^2}{\|\boldsymbol{\gamma}\|_2^2} \tag{11}$$

where T denotes the number of samples, and $\eta_{\theta^{(k)}}^{p^{(k)}}$ is the soft-thresholding function combined with the support selection scheme

$$\eta_{\theta_{(k)}}^{p^{(k)}}(\boldsymbol{\gamma}^k)=\begin{cases}\boldsymbol{\gamma}^k, & i\in\mathcal{S}^{p^{(k)}}(\boldsymbol{\gamma}^k)\\ \eta_{\text{st}}(\boldsymbol{\gamma}^k,\theta_{(k)}), & i\notin\mathcal{S}^{p^{(k)}}(\boldsymbol{\gamma}^k)\end{cases} \tag{12}$$

where $\mathcal{S}^{p^{(k)}}(\boldsymbol{\gamma}^k)$ contains the entries with the $\rho^{(k)}$ largest magnitudes. $\mathbf{W}$ denotes the optimized weight matrix determined with the minimum coherence criterion, which is defined as follows:

$$\begin{aligned}\hat{\mathbf{W}}&=\arg\min_{\mathbf{W}}\mu(\mathbf{W},\mathbf{R})\\&=\arg\min_{\mathbf{W}}\inf_{\mathbf{W}\in\mathbb{C}_{N\times L}}\max_{i\neq j}\mathbf{W}_{:,i}^T\mathbf{R}_{:,j}\\&\quad\text{s.t. }\forall i\in\{1,\ldots,L\}:\mathbf{W}_{:,i}^T\mathbf{R}_{:,i}=1.\end{aligned} \tag{13}$$

Rigorous proof of the convergence and recovery upper and lower bound of HyperLISTA can be found in [18]. An efficient numerical algorithm to calculate the optimized weights is discussed in the Appendix.

Inspired by the outstanding efficiency and performance demonstrated in [18], we consider that HyperLISTA should have great potential in our high-dimensional D-TomoSAR inversion. However, through experiments, we discovered a drawback of HyperLISTA when applied to TomoSAR. Similar to most thresholding algorithms, HyperLISTA suffers from an inherent limitation caused by the global thresholding scheme. Precisely, in the signal projection process for identifying the presence of a dictionary atom within the signal, the selection of an appropriate threshold is of utmost importance. The threshold should be chosen carefully to account for both strong and weak spikes in the reflectivity profile. By selecting a well-suited threshold, the signal projection can distinguish between significant spikes and noise, enabling an accurate identification of dictionary atoms within the signal. However, when utilizing HyperLISTA and other methods that employ global thresholding, the task of selecting an optimal threshold becomes exceedingly challenging. The use of a global threshold implies that the same threshold value is applied uniformly across all entries in the signal. This approach may lead to suboptimal results, as a threshold that effectively captures strong spikes might inadvertently suppress weaker but still meaningful spikes in the reflectivity profile. Consequently, we usually need to choose a relatively small c_1 to have a small threshold so that we can maintain some small spikes caused by reflection from dark scatterers. Otherwise, the information of dark scatterers would be discarded in the thresholding step layer by layer. However, the use of a small threshold brings about two main problems. First, the convergence would be considerably slow. Second, small thresholds yield solutions that are not sparse enough.

To cope with the aforementioned issue and better leverage the power of HyperLISTA in our application, we propose HyperLISTA-ABT, which is an improvement of the original HyperLISTA by incorporation of an ABT (ABT) scheme that explores a local thresholding strategy. The advantages of HyperLISTA-ABT are threefold. First, it conducts the thresholding in each local block, thus allowing for a more refined

thresholding process and possibly retaining weak expressions of reflections from dark scatterers. Then, it becomes possible to better capture the diverse range of spike magnitudes encountered in the signal, enhancing the accuracy and reliability of the reflectivity profile characterization. Second, HyperLISA-ABT has been shown to be more efficient since it updates only one block of variables at each time instead of updating all the variables together. Therefore, HyperLISTA-ABT has been found to be more appropriate for our large-scale and high-dimensional application. Last but not the least, HyperLISTA-ABT reduces the blocksize layerwise and contributes to a better fine-focusing ability.

According to [20] and [21], the update rules of HyperLISTA-ABT after applying block coordinate techniques can be written as follows:

$$\boldsymbol{\gamma}_{i_p}^{k+1} = \eta_{\theta_{i_p}^k}\left(\boldsymbol{\gamma}_{i_p}^k + \boldsymbol{W}_{i_p}^T\left(\boldsymbol{y} - \boldsymbol{R}_{i_p}\boldsymbol{\gamma}_{i_p}^k\right) + \beta_{i_p}^k\left(\boldsymbol{\gamma}_{i_p}^k - \boldsymbol{\gamma}_{i_p}^{k-1}\right)\right) \tag{14}$$

where i_p is the index of the updated block. To clarify, in HyperLISTA-ABT, we remove the support selection scheme and just use the conventional soft-thresholding function. The threshold $\theta^{(k)}$ and the factor $\beta^{(k)}$ are determined for each block as well

$$\theta_{i_p}^k = h_1\left\|\boldsymbol{R}_{i_p}^+\left(\boldsymbol{R}_{i_p}\boldsymbol{\gamma}_{i_p}^k - \boldsymbol{g}\right)\right\|_1 \tag{15}$$

$$\beta_{i_p}^k = h_2\left\|\boldsymbol{\gamma}_{i_p}^k\right\|_0 \tag{16}$$

where $h_1 > 0$, $h_2 > 0$, and $h_3 \in (0, 1)$ are the three hyperparameters. Notably, h_3 is a latent hyperparameter and plays a crucial role in controlling the blocksize despite not explicitly appearing in the formula. In our application, we usually initialize the blocksize according to the grid number within half of the Rayleigh resolution. The block is chosen with a random variants scheme where i_p follows the probability distribution given by

$$P_{i_p} = \frac{L_{i_p}}{\sum_{j=1}^{J} L_{i_p}}, \quad i_p = 1, \ldots, J \tag{17}$$

where J is the number of blocks and $L_{i_p} = ||R_{i_p}^T R_{i_p}||$. All the hyperparameters h_1, h_2, and h_3 can be selected using the same grid search method as in HyperLISTA.

With the blockwise thresholding scheme, local features and weak expressions can be possibly retained. This is due to the fact that many elements of the entries are not strictly driven to zero but to some extremely small value, thus making the output not strictly sparse. Therefore, a postprocessing is usually required to clean the output and make it sparse. The framework of the proposed HyperLISTA-ABT is summarized in Algorithm 1.

Algorithm 1 Summary of the Proposed Algorithm

Generate steering matrix R for given baselines
  Analytic weight optimization **W** according to Eq. (13)
Tuning of hyperparameters
  Simulate ground truth of reflectivity profile $\boldsymbol{\gamma}$ [12]
  Simulate noise-free SAR acquisitions $\mathbf{g} = \mathbf{R}\boldsymbol{\gamma}$
  Grid search to determine the hyperparameters by minimizing NMSE over simulated data
Inference
  Init: $\boldsymbol{\gamma}^0 = \mathbf{R}^H\mathbf{g}$ and blocksize B_1
  for $k = 1, 2, \ldots, K$ **do**
    Determine the number of blocks J_k based on the blocksize B_k
    for $i_p = 1, 2, \ldots, J_k$ **do**

$$\boldsymbol{\gamma}_{i_p}^{k+1} = \eta_{\theta_{i_p}^{(k)}}(\boldsymbol{\gamma}_{i_p}^k + \boldsymbol{W}_{i_p}^T(\boldsymbol{y} - \boldsymbol{R}_{i_p}\boldsymbol{\gamma}_{i_p}^k) + \beta_{i_p}^{(k)}(\boldsymbol{\gamma}_{i_p}^k - \boldsymbol{\gamma}_{i_p}^{k-1}))$$

$$\theta_{i_p}^k = h_1\left\|\boldsymbol{R}_{i_p}^+\left(\boldsymbol{R}_{i_p}\boldsymbol{\gamma}_{i_p}^k - \boldsymbol{g}\right)\right\|_1$$

$$\beta_{i_p}^{(k)} = c_2\left\|\boldsymbol{\gamma}_{i_p}^k\right\|_0$$

    end for
    Update blocksize with $B_{k+1} = h_3 \cdot B_k$
  end for
Output clean-up
Model order selection and final estimation

IV. Simulations

To demonstrate the improvement of the proposed HyperLISTA-ABT to the original HyperLISTA and compare it to the state-of-the-art CS-based and deep learning-based methods, we first conducted experiments based on TomoSAR inversion using simulated data. Since the existing deep learning-based algorithms are not feasible to use with D-TomoSAR cases as explained in Section I, we only focused on TomoSAR inversion for 3-D reconstruction in the simulation.

A. Simulation Setup

In the simulation, we conducted a well-known TomoSAR benchmark test [1], [7] using the same simulation settings as used in [12] and [13]. Specifically, we simulated an interferometric stack containing 25 baselines that are regularly distributed in the range of −135 to 135 m and a two-scatterer mixture in each resolution cell, which is also called double scatterers, meaning that two scatterers are overlaid in the elevation direction in a single pixel. We used the effective detection rate, which is able to simultaneously reflect the super-resolution power and elevation estimation accuracy, to fairly evaluate the performance. Detailed definition of the effective detection rate can be found in [12] and [13].

B. Performance Improvement Compared to the Original HyperLISTA

The first experiment set out to study the performance improvement of HyperLISTA-ABT compared to the original HyperLISTA. In this experiment, the overlaid double scatterers were simulated to have identical scattering phase varying amplitude ratios. The two algorithms were set to have

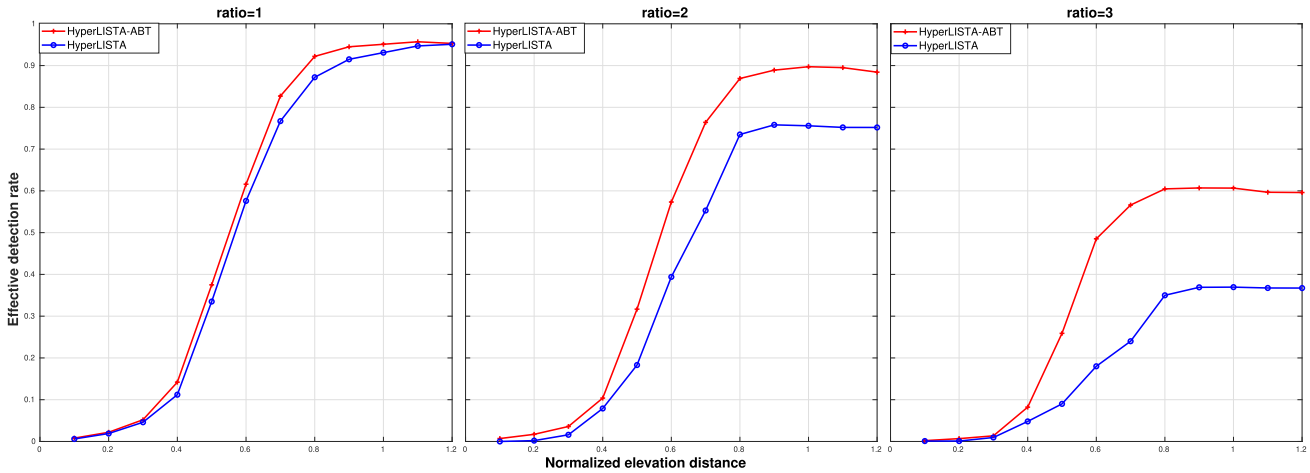


Fig. 3. Effective detection rate of HyperLISTA-ABT and the original HyperLISTA with respect to the normalized elevation distance at different amplitude ratios. The overlaid double scatterers were set to have an identical phase and the SNR level was 6 dB. HyperLISTA-ABT significantly outperformed HyperLISTA at high amplitude ratios between the scatterers.

15 layers, which is a typical number for a LISTA network and its variants. Fig. 3 shows the effective detection rate of HyperLISTA-ABT and the original HyperLISTA as a function of the normalized elevation distance between the simulated double scatterers at 6 dB SNR at different amplitude ratios. Regarding double scatterers, the SNR level and the amplitude ratio are defined as follows.

1) *SNR Level:* For double scatterers, the SNR level is defined such that each individual scatterer is affected by noise corresponding to the given SNR. Both scatterers are set to the same SNR.
2) *Amplitude Ratio:* It refers to the ratio of the amplitudes of the signals from two scatterers. Assuming that a scatterer is represented as $A \cdot e^{i \cdot \phi}$, where A and ϕ represent the amplitude and phase of a deterministic scatterer, respectively, the amplitude ratio can be formally defined as $r_{\text{amp}} = (A_1/A_2)$.

The results demonstrate that HyperLISTA-ABT achieved a significantly higher effective detection rate than the original HyperLISTA. Both algorithms (in fact, all other methods) experience performance degradation with respect to an increase in amplitude ratio. This is attributed to two main factors. First, dark scatterers experience a large bias in their elevation estimates at high amplitude ratios due to their elevation estimates approaching the more prominent ones. Consequently, many detections of double scatterers will not be recognized as effective due to the large elevation estimation bias. Second, the energy of dark scatterers is close to the noise level at high amplitude ratios. This makes it particularly challenging for HyperLISTA, which employs a global thresholding scheme, to detect the local features of dark scatterers. Further elaborating, when high-intensity scatterers are present in the signal, their strong energy can overshadow the low-energy regions where dark scatterers are located. This overshadowing effect can lead to the suppression or even annihilation of the weaker expressions associated with the dark scatterers. Consequently, the presence of these strong intensity scatterers can mask or obscure the signals originating from the dark scatterers, making their detection and characterization challenging. In contrast, HyperLISTA-ABT conducts thresholding in each local block, which can allow retaining local information and, thus, it can detect dark scatterers. This results in a higher effective detection rate at high amplitude ratios.

C. Comparison With the State-of-the-Art Algorithms

In this section, we compared HyperLISTA-ABT to other state-of-the-art algorithms for further evaluation, which are deep learning-based algorithms γ-Net [12] and CV-SMGUs [13], as well as the traditional CS-based method SL1MMER [22] with second-order optimization. To highlight the super-resolution ability of these methods, we also involved a conventional spectral estimator SVD-Wiener [1] as a baseline in the comparison.

The comparison was first based on the effective detection rate. Two different scenarios were taken into consideration: SNR $\in \{0, 6\}$ dB, which represents a noisy case, and a regular case with a typical SNR level in a high-resolution spaceborne SAR image. The comparison results are demonstrated in Fig. 4. At each discrete normalized elevation distance, 0.2 million Monte Carlo trials with an identical phase and amplitude, which represents the worst case [22] in TomoSAR inversion, were simulated. The deep learning-based algorithms γ-Net and CV-SMGUs were built with 12 and 6 hidden layers, respectively. The training followed the same training strategy introduced in [12] and [13] and was carried out using a single NVIDIA RTX2080 GPU. For HyperLISTA-ABT, the training involved analytical weight optimization and determining the hyperparameters via the grid search method. The number of iterations in HyperLISTA-ABT was set as 15.

From the comparison results, we can see that all the methods except the conventional spectral estimator SVD-Wiener showed a great super-resolution power. The proposed HyperLISTA-ABT delivered almost the same super-resolution ability as γ-Net and approached the performance of

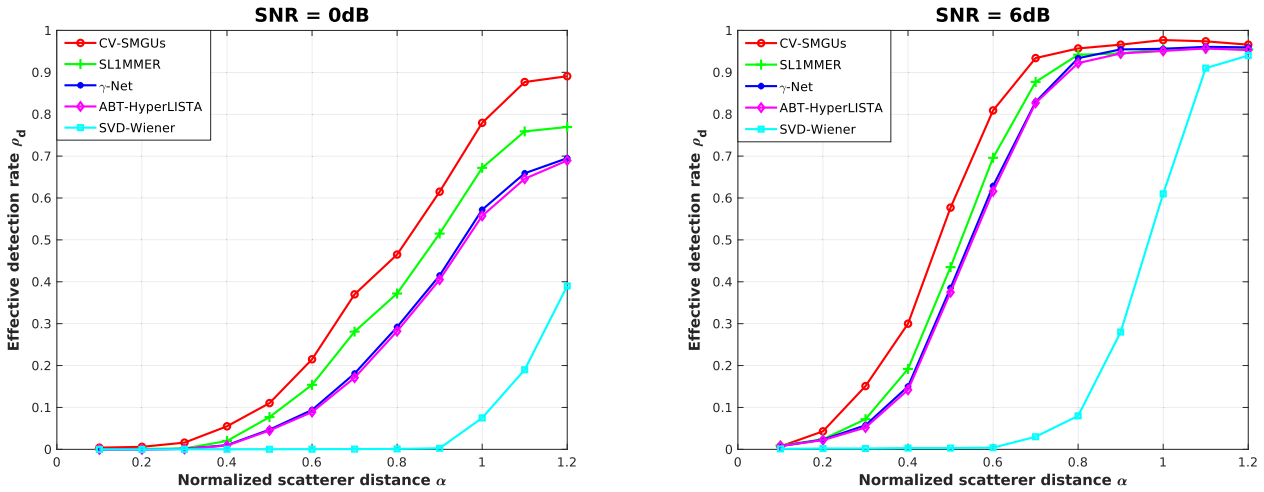


Fig. 4. Detection rate P_d as a function of the normalized elevation distance between the simulated facade and ground with SNR $=$ 0 dB and 6 dB, $N = 25$, and phase difference $\triangle\phi = 0$ (worst case) under 0.2 million Monte Carlo trials.

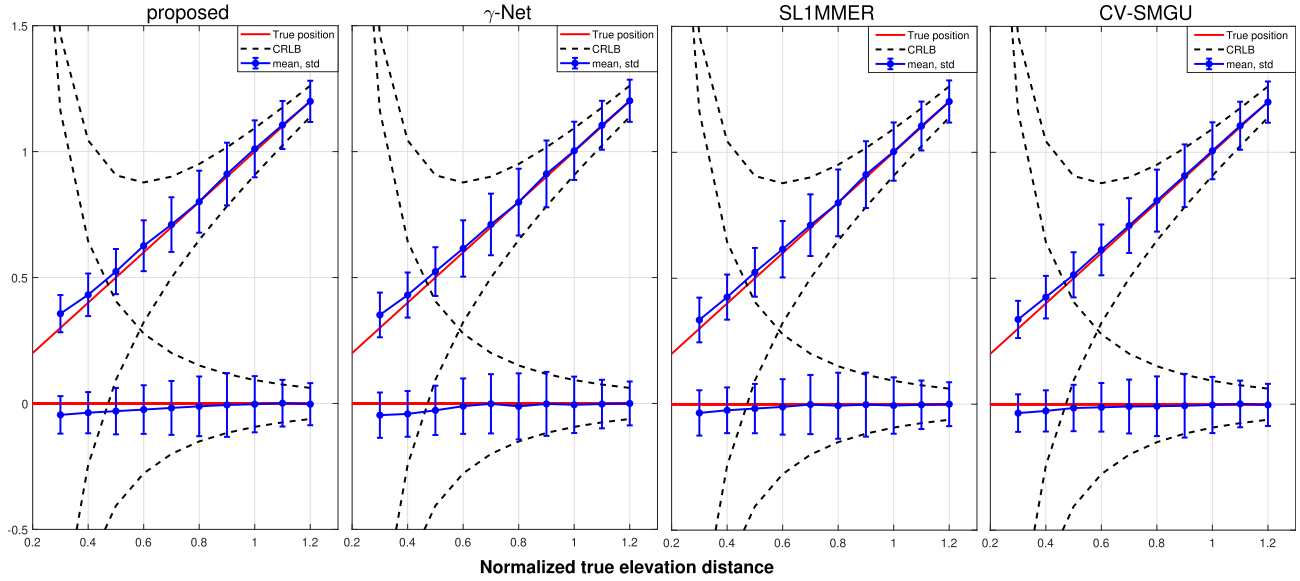


Fig. 5. Estimated elevation of simulated facade and ground. The SNR level is set as 6 dB. Each dot has the sample mean of all estimates as its y value and the corresponding standard deviation as error bar. The red line segments represent the true elevation of the simulated facade and ground. The dashed curves denote the true elevation $\pm 1\times$ CRLB normalized with respect to the Rayleigh resolution.

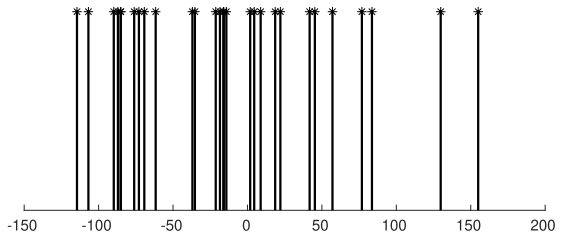


Fig. 6. Effective baselines of the 29 TerraSAR-X high-resolution spotlight images.

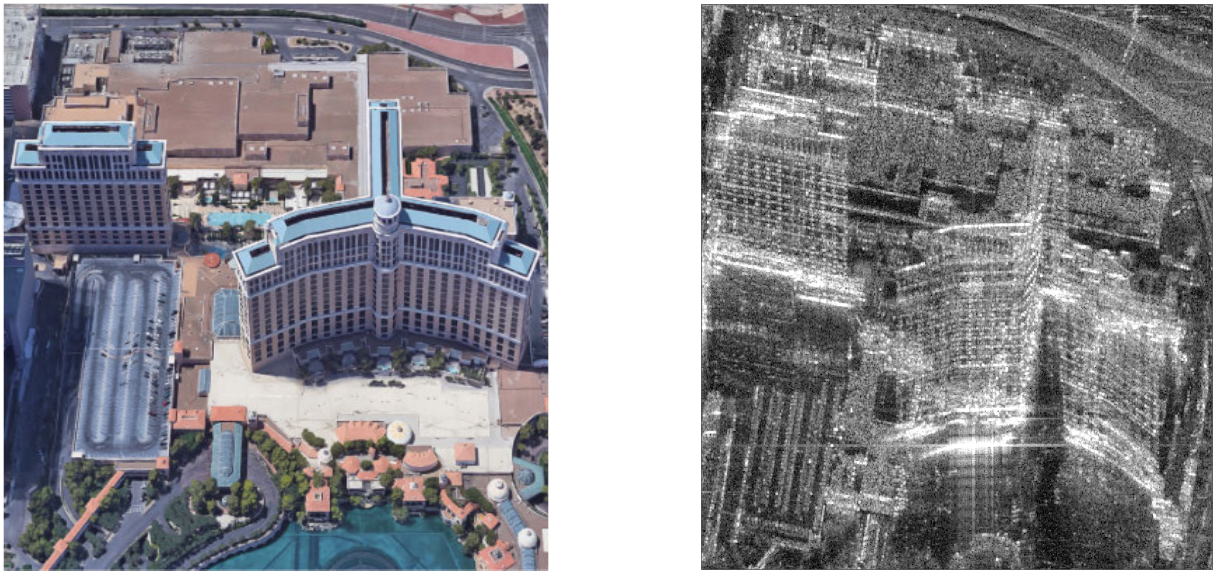
SL1MMER in both scenarios. In addition, we plotted the elevation estimates of the simulated facade and ground in Fig. 5 with respect to the normalized true elevation distance to demonstrate the elevation estimation accuracy of the tested super-resolution algorithms. As depicted in Fig. 5, all evaluated algorithms demonstrate comparable accuracy in elevation estimation. This consistency across various algorithms indicates that our effective detection constraint effectively guarantees elevation estimation accuracy. However, when focusing solely on the effective detection rate, it was challenging to proclaim a clear advantage of the proposed HyperLISTA-ABT method over the existing state-of-the-art approaches. In fact, when comparing it to CV-SMGUs, we could observe a slight underperformance. However, all the state-of-the-art methods come with a relatively high computational cost. Both $\boldsymbol{\gamma}$-Net and CV-SMGU require tailored training according to the baseline distribution of the stack. SL1MMER is a model-based algorithm, thus needs no training, yet requires significantly computational time for solving the $L1$-norm minimization.

We tested and recorded the time consumption of different algorithms for processing the 0.2 million Monte Carlo trials as well as the requirements for training data. The results are

TABLE I

COMPARISON OF THE NUMBER OF REQUIRED TRAINING SAMPLES AND TIME CONSUMPTION FOR PROCESSING 0.2 MILLION MONTE CARLO TRIALS WITH EACH ALGORITHM. THE TRAINING TIME OF HYPERLISTA-ABT INDICATES THE COMBINED DURATION OF BOTH THE ANALYTIC WEIGHT OPTIMIZATION PROCESS AND THE TUNING OF HYPERPARAMETERS. IT PROVIDES A MEASURE OF THE OVERALL TIME REQUIRED FOR THESE ESSENTIAL STEPS

Algorithm	number of training samples	training time	inference time	total time consumption	transferability
CV-SMGUs	4 million	$\approx$ 10 hours	$\approx$ 0.25 h	$\approx$ 10 h	low
γ-Net	3 million	$\approx$ 8 hours	$\approx$ 0.2 h	$\approx$ 8 h	low
SL1MMER	-	-	$\approx$ 20 h	$\approx$ 20 h	high
HyperLISTA-ABT	-	$\approx$ 0.5 hour	$\approx$ 0.25 h	$\approx$ 1 h	medium



(a) (b)

Fig. 7. Test site. (a) Optical image from Google Earth. (b) SAR mean intensity image.

summarized in Table I. To clarify, all inferences were conducted using a local CPU for a fair comparison. As can be seen in Table I, it took about 10 h for the deep learning-based algorithms to process 0.2 million Monte Carlo trials, which was predominantly the training time. In addition, a large amount of training samples was essential as well. For SL1MMER, it took about 20 h for the processing since the iterative second-order optimization is computationally expensive. Further inspecting the table, we can see that HyperLISTA-ABT showed similar efficiency in the inference as the other deep learning-based algorithms. However, HyperLISTA-ABT required no training data and it took much less time for the training. In total, HyperLISTA-ABT speeded up the processing by about one order of magnitude compared to the other algorithms tested in the experiment.

Upon evaluating the performance and efficiency, it was observed that HyperLISTA-ABT achieved comparable performance to the existing state-of-the-art methods while significantly improving the computational efficiency by approximately one order of magnitude. This is especially advantageous in the multicomponent D-TomoSAR case. The application of the aforementioned deep learning-based algorithms and SL1MMER is very limited in the D-TomoSAR case due to the need for time-consuming model training and the heavy computational expense. On the contrary, the application of HyperLISTA-ABT can be easily extended to computationally efficient D-TomoSAR processing. Therefore, HyperLISTA-ABT is a more applicable approach for the large-scale processing of real data.

Furthermore, HyperLISTA-ABT demonstrates superior transferability compared to deep learning-based algorithms. Deep learning models are typically trained to fit specific baseline configurations, such as a fixed number of SAR acquisitions and a specific baseline distribution. While they may exhibit satisfactory generalizability to small baseline discrepancies [12], [13], directly applying a trained deep learning model to a new data stack with a different number of acquisitions or a completely different baseline distribution is not feasible. In such cases, time-consuming retraining of the model becomes necessary, resulting in low transferability.

In contrast, HyperLISTA-ABT offers better transferability. Although it requires analytical optimization of the weight matrix for each new data stack, the efficiency of the analytical optimization process allows for scalability and improved transferability. This finding highlights the potential of HyperLISTA-ABT in enabling global urban mapping using TomoSAR, as it can be effectively applied to diverse data

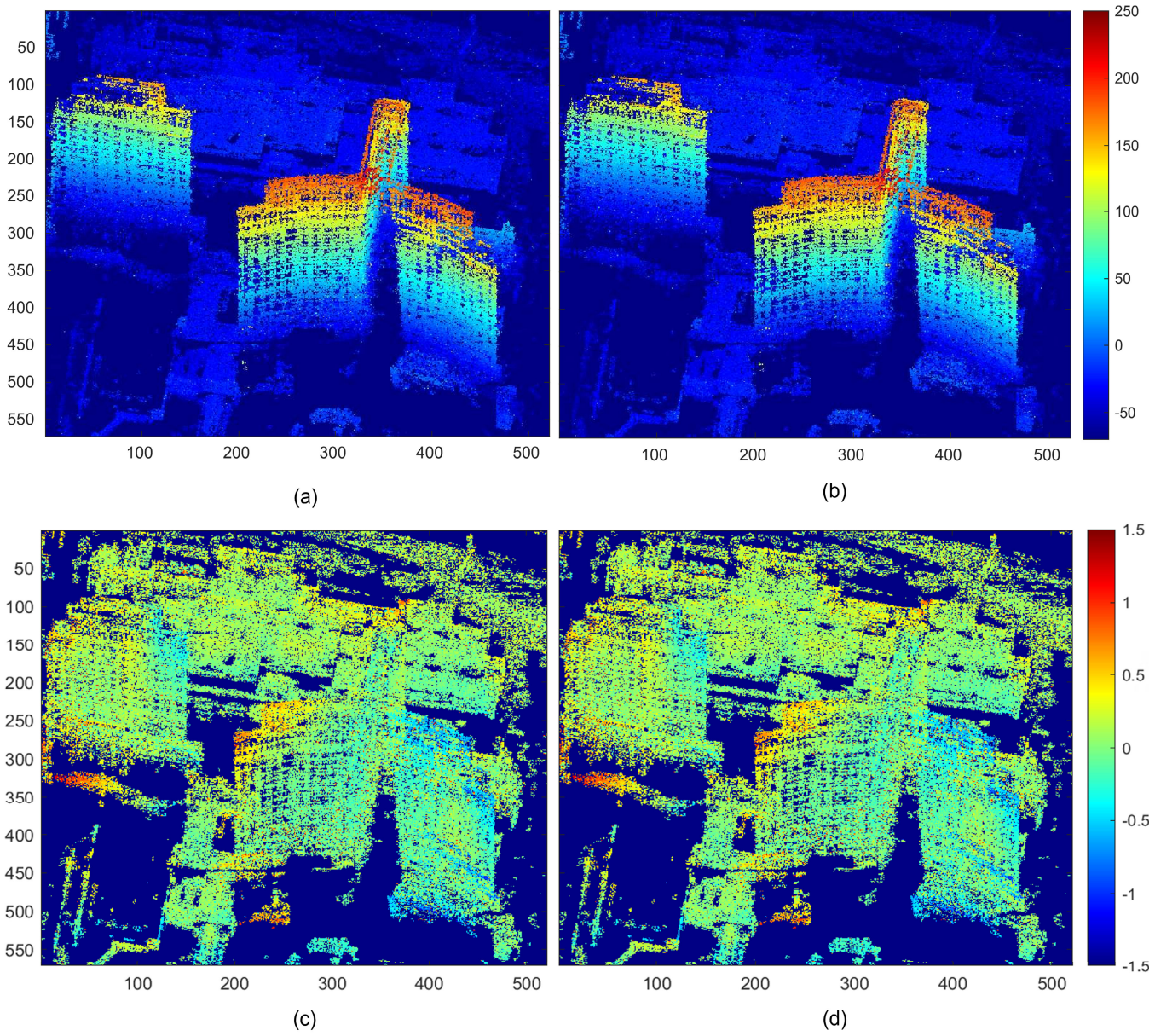


Fig. 8. Color-coded reconstruction results of the test site. (a) Elevation estimates using HyperLISTA-ABT in meters, (b) elevation estimates using SL1MMER in meters, (c) estimated amplitude of seasonal motion using HyperLISTA-ABT in centimeters, and (d) estimated amplitude of seasonal motion using SL1MMER in centimeters.

stacks with varying acquisition configurations and baseline distributions.

V. Real Data Experiment

A. Bellagio Hotel

In this real data experiment, due to the fact that there was no available ground truth, we purposely used the same data as in [23] so that we can compare our results to the results obtained with SL1MMER. The data stack was composed of 29 TerraSAR-X high-resolution spotlight images covering the Bellagio Hotel in Las Vegas, whose baseline distribution is illustrated in Fig. 6. The slant range resolution was 0.6 m and the azimuth resolution was 1.1 m. The elevation aperture size of about 270 m resulted in the inherent elevation resolution ρ_s to be about 40 m, i.e., approximately 20 m resolution in height since the incidence angle here was 31.8°. An optical image and the SAR mean intensity image of the test site are shown in Fig. 7.

As for the D-TomoSAR system model, a time wrap operation assuming only sinusoidal seasonal motion was adopted as in [16] because no long-term linear motion was observed during the acquisition period of the test area.

In Fig. 8, we compare the estimated elevation and amplitude of the seasonal motion of the detected single scatterers and the top layer of the detected double scatterers. From Fig. 8(a), we can see a smooth gradation of the elevation estimates from the building bottom top, which suggests a reasonable elevation estimation by HyperLISTA-ABT. Moreover, we can see that there is no significant difference

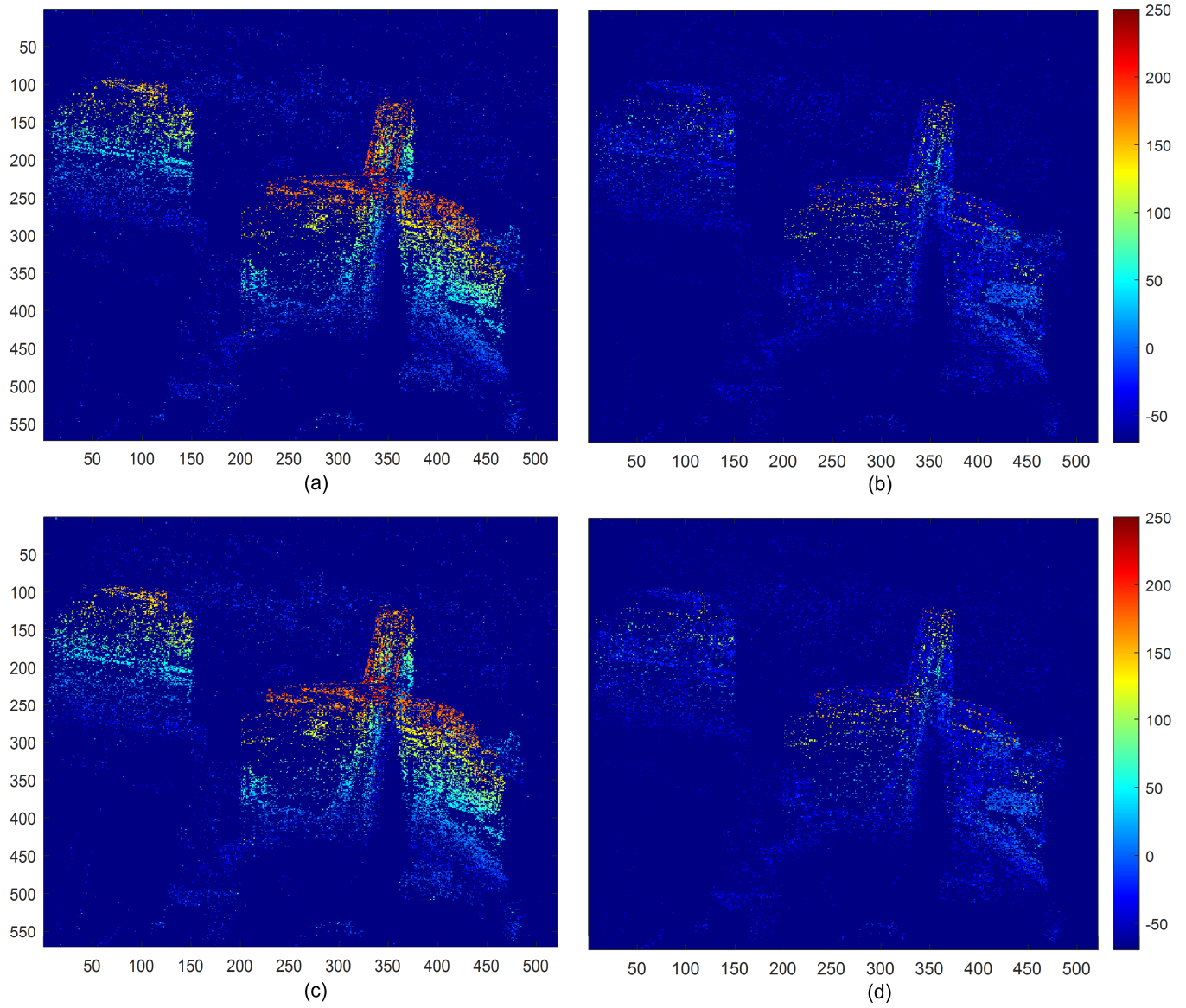


Fig. 9. Color -coded elevation estimates of the top and bottom layers of the detected double scatterers using HyperLISTA-ABT (a) and (b) and SL1MMER (c) and (d). From left to right are the top and bottom layers of the detected double scatterers. The top layer is mostly caused by reflections from the building roof and façade, while the bottom layer is mostly caused by reflections from low infrastructures and the ground.

between the results of HyperLISTA-ABT and SL1MMER, implying that HyperLISTA-ABT had similar performance to SL1MMER. In addition, Fig. 9 shows the layover separation ability of HyperLISTA-ABT. As can be seen, the two layers of double scatterers were detected and separated by HyperLISTA-ABT. The top layer was mainly caused by signals from the roof and facade of the high-rise buildings while the bottom layer was caused by signals from the ground structures. By comparing the reconstruction results of HyperLISTA-ABT and SL1MMER, we can see that HyperLISTA-ABT achieves a remarkably similar super-resolution power to SL1MMER.

We also conducted some numerical comparisons of both algorithms. First, we compared the percentage of pixels detected as zero, one, and two scatterers by both algorithms. A comparison of the number of scatterers detected by the proposed HyperLISTA-ABT and SL1MMER is demonstrated in Fig. 10 and the detailed statistics is listed in Table II. Compared to SL1MMER, we found that HyperLISTA-ABT detected more pixels as coherent scatterers. This does not necessarily mean that HyperLISTA-ABT had a better detection ability since there was no ground truth. We believe that HyperLISTA-ABT detected more scatterers because HyperLISTA-ABT tends to maintain weak signals, which could be reflections of dark scatterers but also outliers caused by noise interference. The false detection of noise as coherent scatterers causes a speckle-like noise in the reconstruction result. Model order selection and postprocessing techniques like spatial filtering can further mitigate such outliers.

For further evaluation, we compared the elevation estimates differences of scatterers detected by both algorithms. The

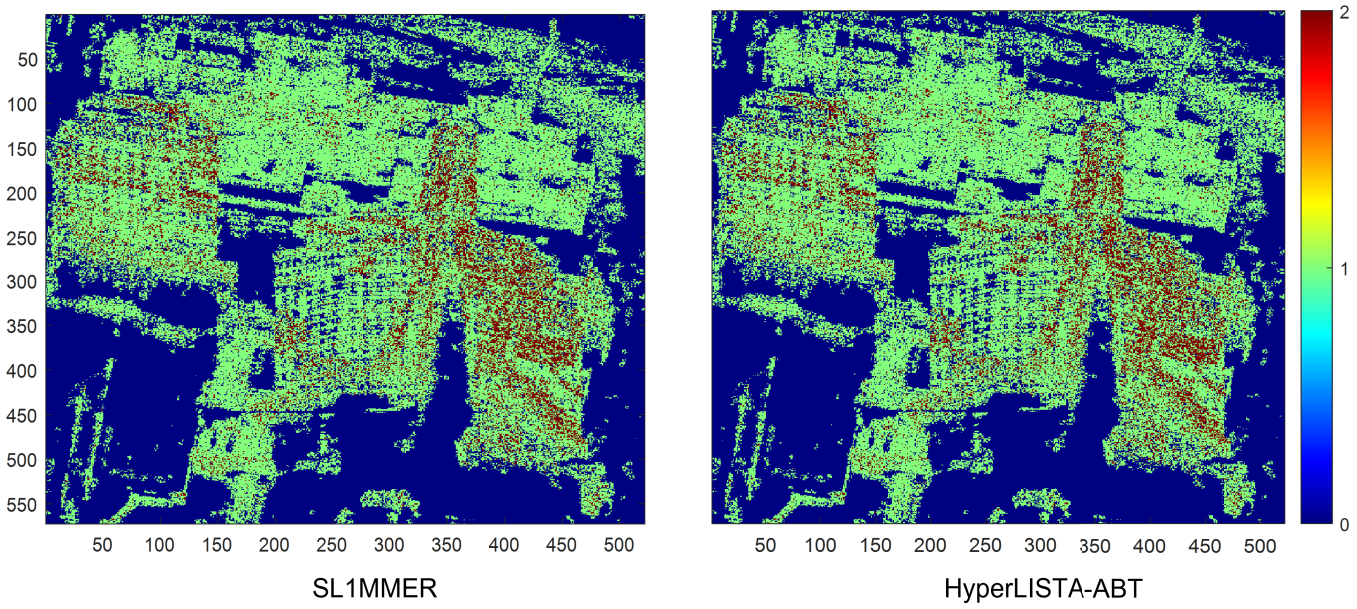


Fig. 10. Map of the number of scatterers detected by HyperLISTA-ABT and SL1MMER.

TABLE II
PERCENTAGE OF SCATTERERS DETECTION FOR THE TWO ALGORITHMS

Algorithm	Percentage of detection as		
	0 scatterer	1 scatterer	2 scatterers
HyperLISTA-ABT	48.48 %	44.09 %	7.43 %
SL1MMER	49.41 %	43.63 %	6.96 %

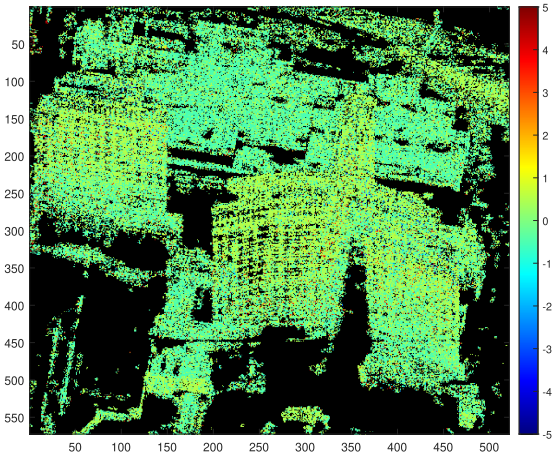


Fig. 11. Demonstration of elevation estimates differences in meters between HyperLISTA-ABT and SL1MMER.

differences of the elevation estimates are demonstrated in Fig. 11 and a histogram of the elevation estimates differences is shown in Fig. 12. Since the elevation coordinates were discretized on a 0.5 m grid the intervals of the elevation estimates differences follow the 0.5 m increments. It can be observed that most of the elevation estimates differences were within 1 m. This observation indicates that both algorithms yielded comparable results in terms of elevation estimation,

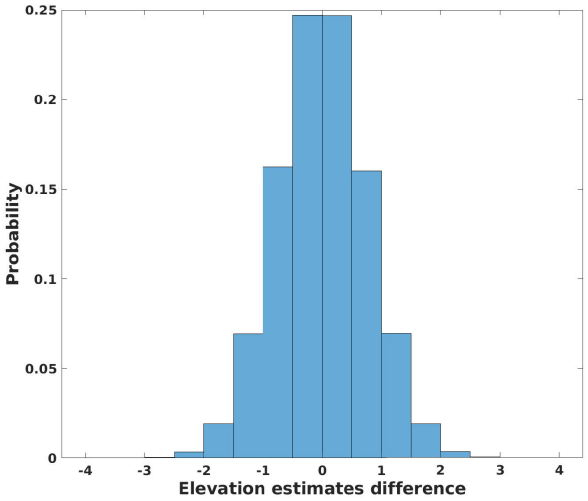


Fig. 12. Histogram of elevation estimates differences between HyperLISTA-ABT and SL1MMER.

instilling confidence in their reliability and reasonableness. Furthermore, this similarity in estimation accuracy suggests that HyperLISTA-ABT performed on par with SL1MMER. Moreover, it is worth mentioning that it took more than three weeks for SL1MMER to finish the D-TomoSAR processing over the test site, whereas it only took several hours for HyperLISTA-ABT to complete the processing.

B. Large Demonstration

In this section, we applied HyperLISTA-ABT to TerraSAR-X high-resolution spotlight data over a large area surrounding the convention center in Las Vegas. The stack was composed of 29 acquisitions covering a time period from July 2009 to June 2010, during which the test

Fig. 13. Demonstration of the large test area. (a) Optical image from Google Earth. (b) SAR mean intensity map in decibel. The red box in (b) indicates the area undergoing subsidence.

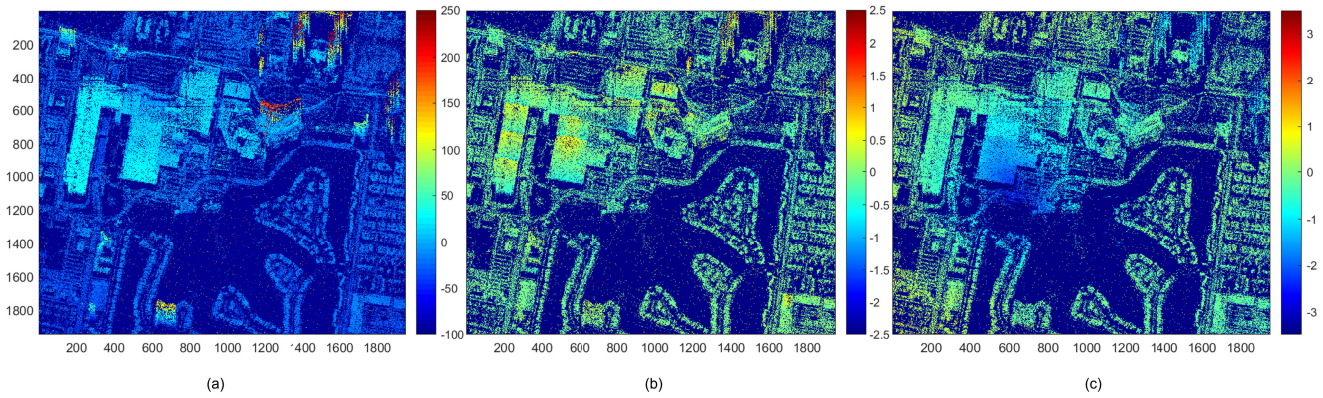


Fig. 14. Demonstration of color-coded elevation estimates and estimated amplitude of multicomponent motion. (a) Elevation estimates in meters, (b) estimated amplitude of seasonal motion in centimeters, and (c) estimated amplitude of linear motion in centimeters/year.

area was undergoing a pronounced subsidence centered at the convention center. Therefore, the acquisitions were characterized by a multicomponent nonlinear motion combining linear and thermal dilation-induced seasonal motion. Fig. 13 shows an optical image and the SAR mean intensity map. The red box indicates the "epicenter" undergoing subsidence around the convention center.

Fig. 14 illustrates the reconstructed elevation estimates as well as the estimated amplitude maps of the two different motions. As we can see from the surface model generated from the elevation estimates in Fig. 14(a), we can capture the shapes of individual buildings and the surrounding infrastructures, like roads, at a detailed level. In addition, Fig. 14(b) shows that clear deformation caused by seasonal motion can be observed in the metallic building structures since they were affected by thermal dilation more seriously compared to the surrounding infrastructures. Furthermore, as illustrated in Fig. 14(c), it could be observed that the magnitude of the linear subsidence increased as the scatterer gets closer to the "epicenter." These results are consistent with the fact, thus validating the effectiveness of HyperLISTA-ABT for multicomponent nonlinear motion estimation and giving confidence that HyperLISTA-ABT can be applied in large-scale D-TomoSAR processing.

VI. Discussion

A. Key Differencing Features Compared to ALISTA

The proposed HyperLISTA-ABT differs from ALISTA [17] in several key aspects, summarized as follows.

1) *Enhance Convergence:* HyperLISTA-ABT incorporates Polyak's heavy ball method [24], introducing a momentum component that creates layerwise skip connections in the model. This innovation notably accelerates the convergence rate. Consequently, HyperLISTA-ABT achieves faster linear convergence compared to ALISTA under identical conditions.

2) *Instance-Adaptive Parameter Design:* The model utilizes an instance-adaptive parameter design [18], simplifying it to three hyperparameters, which can be efficiently tuned manually. Therefore, HyperLISTA-ABT eliminates the need for extensive GPU memory usage for storing the huge model.
3) *Adaptive Blockwise Thresholding:* HyperLISTA-ABT allows for a more refined thresholding process, thus enhancing the detection and preservation of weak signals from less reflective scatterers. Moreover, it updates a single block of variables at a time rather than simultaneously updating all variables, leading to greater efficiency compared to ALISTA.

B. Description of the Grid Search Method for Determining the Hyperparameters

Grid search [25] is an established method in the realm of machine learning for hyperparameter optimization. It systematically explores a specified subset of hyperparameters, thereby facilitating the identification of the most efficacious combination to enhance model performance. The specific implementation of grid search for HyperLISTA-ABT is as follows.

1) *Range Specification:* We established ranges for hyperparameters h_1, h_2, and h_3. Based on our experience, the initial ranges are chosen as $(0, 0.1)$ for both h_1 and h_2, and $(0.9, 1)$ for h_3.
2) *Coarse Grid Generation:* For each hyperparameter (h_1, h_2, h_3), we generated ten coarse grids, labeled as x, y, and z, respectively.
3) *Model Evaluation:* For each combination of hyperparameters in grids, we calculated the NMSE over the validation set.
4) *Search Range Refinement:* The grid combination (x_{s1}, y_{s2}, z_{s3}) yielding the lowest NMSE was selected. We then refined our search to the adjacent grids, defining new ranges as (x_{s1-1}, x_{s1+1}), (y_{s2-1}, y_{s2+1}), and (z_{s3-1}, z_{s3+1}) for h_1, h_2, and h_3, respectively. In cases where the selected grid was at the boundary (first or last), the range was adjusted accordingly. For instance, if $s1 = 1$, the new range for h_1 becomes (x_{s1}, x_{s1+1}) instead of (x_{s1-1}, x_{s1+1}).
5) *Iteration:* We repeated steps 2 to 4 until we achieved convergence.

VII. Conclusion

This article proposes HyperLISTA-ABT to address the gap in applying deep neural networks for solving D-TomoSAR inversion. Unlike traditional methods that learn weights directly from data, HyperLISTA-ABT computes the weights with an analytical optimization technique by minimizing generalized mutual coherence. Additionally, HyperLISTA-ABT introduces an ABT scheme that applies block coordinate techniques to accelerate the algorithm. Moreover, it conducts thresholding in local blocks to retain weak expressions of reflection from dark scatterers and considers more local features. Laboratory experiments for 3-D reconstruction confirmed the efficiency of HyperLISTA-ABT in estimation. Moreover, tests on real data over a large area demonstrated that HyperLISTA-ABT can reconstruct high-quality 4-D point clouds, making it an efficient and accurate algorithm for future large-scale or even global D-TomoSAR processing.

Algorithm 2 Efficient Algorithm for Analytical Weight Optimization

Input: the steering matrix $\mathbf{R}$
Init: $\mathbf{D} = \mathbf{R}$, $\mathbf{G} = \mathbf{I}$, $\zeta = \alpha = 0.1$
for iter $= 1, 2, \ldots$ until **convergence do**
  update $\mathbf{D}$ with (20)
  update $\mathbf{G}$ with (21)
  Compute $f_1 = \left\| \boldsymbol{D}^H \boldsymbol{D} - \boldsymbol{I} \right\|_F^2$
  Compute $f_2 = \left\| (\boldsymbol{G}\boldsymbol{R})^H \boldsymbol{G}\boldsymbol{A} - \boldsymbol{I} \right\|_F^2$
  if two consecutive f_1s are close enough **then**
    $\zeta = 0.1\zeta$
    $\alpha = 0.1\alpha$
    if f_1 and f_2 are close enough **then**
      break
    end
  end
end
Output: $\mathbf{W} = \mathbf{G}^H\mathbf{G}\mathbf{R}$

Appendix

A. Efficient Algorithm to Calculate Weight Analytically

As discussed in [17], it is difficult to solve (13) directly and (13) can be reformulated as minimizing the Frobenius norm of $\mathbf{W}^H\mathbf{R}$ over a linear constraint. Defining $\mathbf{W} = \mathbf{G}^H\mathbf{G}\mathbf{R}$ ($\mathbf{G} \in \mathbb{C}^{N\times N}$ is named as the Gram matrix), the minimization of the Frobenius norm reads

$$\min_{G} \left\| R^H \boldsymbol{G}^H \boldsymbol{G}\boldsymbol{R} - \boldsymbol{I} \right\|_F^2, \quad \text{s.t. } \operatorname{diag}\left(\boldsymbol{R}^H \boldsymbol{G}^H \boldsymbol{G}\boldsymbol{R}\right) = \mathbf{1}. \tag{18}$$

However, it is hard to handle the constraint in the above problem (18). As a solution, a matrix $\mathbf{D} = \mathbf{G}\mathbf{R} \in \mathbb{C}^{N\times L}$ is introduced and we use the following method as an alternative:

$$\begin{aligned} &\min_{\boldsymbol{G},\boldsymbol{D}} \left\| \boldsymbol{D}^T \boldsymbol{D} - \boldsymbol{I} \right\|_F^2 + \frac{1}{\alpha} \|\boldsymbol{D} - \boldsymbol{G}\boldsymbol{R}\|_F^2 \\ &\text{s.t. } \operatorname{diag}\left(\boldsymbol{D}^T \boldsymbol{D}\right) = \mathbf{1}. \end{aligned} \tag{19}$$

With a proper $\alpha > 0$, the solution to (19) approximates (18) and we obtain the optimized weights accordingly. The steps for solving the optimization problem (19) are described as follows.

First, $\mathbf{G}$ is fixed and we update $\mathbf{D}$ with the projected gradient descent (PGD)

$$\boldsymbol{D} \leftarrow \mathcal{P}\left(\boldsymbol{D} - \zeta \boldsymbol{D}\left(\boldsymbol{D}^H \boldsymbol{D} - \boldsymbol{I}\right) - \frac{\zeta}{\alpha}(\boldsymbol{D} - \boldsymbol{G}\boldsymbol{R})\right) \tag{20}$$

where $\mathcal{P}$ denotes the projection operator on the constraint $\operatorname{diag}(\boldsymbol{D}^T\boldsymbol{D}) = \mathbf{1}$, so that each column of $\mathbf{D}$ will be normalized, and ζ is the stepsize. Hereafter, we fix D and update the minimizer of G with

$$\boldsymbol{G} \leftarrow \mathbf{D}\mathbf{R}^+ \tag{21}$$

where $\mathbf{R}^{+}$ represents the Moore–Penrose pseudo inverse of the steering matrix $\mathbf{R}$. Then, we repeat the procedure until $\mathbf{D} \approx \mathbf{GR}$. The whole algorithm is summarized in Algorithm 2.

ACKNOWLEDGMENT

The authors acknowledge DLR for providing the TerraSAR-X data via the science program. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.

[2] D. Reale, G. Fornaro, A. Pauciullo, X. Zhu, and R. Bamler, "Tomographic imaging and monitoring of buildings with very high resolution SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 661–665, Jul. 2011.

[3] X. X. Zhu and M. Shahzad, "Facade reconstruction using multiview spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3541–3552, Jun. 2014.

[4] G. Fornaro, F. Lombardini, A. Pauciullo, D. Reale, and F. Viviani, "Tomographic processing of interferometric SAR data: Developments, applications, and future research perspectives," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 41–50, Jul. 2014.

[5] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[6] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[7] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by L_1-norm regularization—The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.

[8] Y. Shi, X. X. Zhu, W. Yin, and R. Bamler, "A fast and accurate basis pursuit denoising algorithm with application to super-resolving tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6148–6158, Oct. 2018.

[9] X. X. Zhu, N. Ge, and M. Shahzad, "Joint sparsity in SAR tomography for urban mapping," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1498–1509, Dec. 2015.

[10] A. Budillon, A. C. Johnsy, G. Schirinzi, and S. Vitale, "SAR tomography based on deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 3625–3628.

[11] J. R. Hershey, J. Le Roux, and F. Weninger, "Deep unfolding: Model-based inspiration of novel deep architectures," 2014, *arXiv:1409.2574*.

[12] K. Qian, Y. Wang, Y. Shi, and X. X. Zhu, "γ-Net: Superresolving SAR tomographic inversion via deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4706116.

[13] K. Qian, Y. Wang, P. Jung, Y. Shi, and X. X. Zhu, "Basis pursuit denoising via recurrent neural network applied to super-resolving SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4710015.

[14] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.

[15] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.

[16] X. X. Zhu and R. Bamler, "Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 735–739, Jul. 2011.

[17] J. Liu, X. Chen, Z. Wang, and W. Yin, "ALISTA: Analytic weights are as good as learned weights in LISTA," in *Proc. Int. Conf. Learn. Represent.*, 2019, pp. 1–13. [Online]. Available: https://openreview.net/forum?id=B1lnzn0ctQ

[18] X. Chen, J. Liu, Z. Wang, and W. Yin, "Hyperparameter tuning is all you need for LISTA," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, pp. 1–12.

[19] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[20] Y. Xu and W. Yin, "Block stochastic gradient iteration for convex and nonconvex optimization," *SIAM J. Optim.*, vol. 25, no. 3, pp. 1686–1716, 2015.

[21] Z. Peng, T. Wu, Y. Xu, M. Yan, and W. Yin, "Coordinate-friendly structures, algorithms and applications," *Ann. Math. Sci. Appl.*, vol. 1, no. 1, pp. 57–119, 2016.

[22] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, Jan. 2012.

[23] X. X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3150–3157, Aug. 2012.

[24] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Comput. Math. Math. Phys.*, vol. 4, no. 5, pp. 1–17, Jan. 1964.

[25] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed. Berlin, Germany: Springer, 2009. [Online]. Available: http://www-stat.stanford.edu/ tibs/ElemStatLearn/

Kun Qian received double B.Sc. degree in remote sensing and information engineering and aerospace engineering and geodesy from Wuhan University, Wuhan, China, and the University of Stuttgart, Stuttgart, Germany, in 2016, and the M.Sc. degree in aerospace engineering and geodesy from the University of Stuttgart in 2018. He is currently pursuing the Ph.D. degree in data science in earth observation, Technical University of Munich, Munich, Germany.

His research interests include data-driven methods and deep unfolding algorithms and their application in multibaseline SAR tomography.

Yuanyuan Wang (Member, IEEE) received the B.Eng. degree (Hons.) in electrical engineering from The Hong Kong Polytechnic University, Hong Kong, China, in 2008, and the M.Sc. and Dr.-Ing. degrees from the Technical University of Munich, Munich, Germany, in 2010 and 2015, respectively.

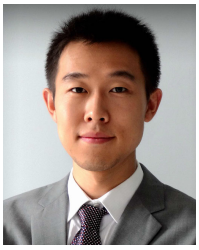
From June to July 2014, he was a Guest Scientist with the Institute of Visual Computing, ETH Zürich, Zürich, Switzerland. He is currently a Guest Professor with the German International AI Future Laboratory (AI4EO), Technical University of Munich. His research interests include optimal and robust parameters estimation in multibaseline InSAR techniques, multisensor fusion algorithms of SAR and optical data, nonlinear optimization with complex numbers, machine learning in SAR, uncertainty quantification and mitigation in machine learning, and high-performance computing for big data.

Dr. Wang was one of the best reviewers of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2016.

Peter Jung (Member, IEEE) received the Dipl.-Phys. degree in high energy physics from DESY Hamburg, Humboldt University, Berlin, Germany, in 2000, and the Dr.-rer.nat (Ph.D.) degree from Weyl–Heisenberg Representations in Communication Theory, Technical University of Berlin (TUB), Berlin, in 2007.

Since 2001, he has been with the Department of Broadband Mobile Communication Networks, Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut (HHI), Berlin, and since 2004 he has been with the Fraunhofer German-Sino Laboratory for Mobile Communications, Berlin. He is currently working under DFG grants in the field of signal processing, information and communication theory, and data science at TUB. He is also a Visiting Professor with TUM, Munich, Germany, and associated with the Munich AI Future Laboratory (AI4EO). He is giving lectures in compressed sensing, estimation theory, and inverse problems. His research interests include the area-compressed sensing, machine learning, time–frequency analysis, dimension reduction, and randomized algorithms.

Dr. Jung is also a member of VDE/ITG.



Yilei Shi (Member, IEEE) received the Dipl.-Ing. degree in mechanical engineering and the Dr.-Ing. degree in signal processing from Technische Universität München (TUM), Munich, Germany, in 2010 and 2019, respectively.

His research interests include fast solver and parallel computing for large-scale problems, high-performance computing and computational intelligence, advanced methods on SAR and InSAR processing, machine learning and deep learning for a variety of data sources, such as SAR, optical images, and medical images, and PDE-related numerical modeling and computing.

Xiao Xiang Zhu (Fellow, IEEE) received the Master (M.Sc.) degree, her doctor of engineering (Dr.-Ing.) degree and her "Habilitation" in the field of signal processing from Technical University of Munich (TUM), Munich, Germany, in 2008, 2011 and 2013, respectively.

She is the Chair Professor for Data Science in Earth Observation at Technical University of Munich (TUM) and was the founding Head of the Department "EO Data Science" at the Remote Sensing Technology Institute, German Aerospace Center (DLR). Since May 2020, she is the PI and director of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond", Munich, Germany. Since October 2020, she also serves as a Director of the Munich Data Science Institute (MDSI), TUM. From 2019 to 2022, Zhu has been a co-coordinator of the Munich Data Science Research School (www.mu-ds.de) and the head of the Helmholtz Artificial Intelligence – Research Field "Aeronautics, Space and Transport". Prof. Zhu was a guest scientist or visiting professor at the Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, the University of Tokyo, Tokyo, Japan and University of California, Los Angeles, United States in 2009, 2014, 2015 and 2016, respectively. She is currently a visiting AI professor at ESA's Phi-lab, Frascati, Italy. Her main research interests are remote sensing and Earth observation, signal processing, machine learning and data science, with their applications in tackling societal grand challenges, e.g. Global Urbanization, UN's SDGs and Climate Change.

Dr. Zhu has been a member of young academy (Junge Akademie/Junges Kolleg) at the Berlin-Brandenburg Academy of Sciences and Humanities and the German National Academy of Sciences Leopoldina and the Bavarian Academy of Sciences and Humanities. She is a Fellow of the Academia Europaea (the Academy of Europe). She serves in the scientific advisory board in several research organizations, among others the German Research Center for Geosciences (GFZ, 2020-2023) and Potsdam Institute for Climate Impact Research (PIK). She is an associate Editor of IEEE Transactions on Geoscience and Remote Sensing, Pattern Recognition and served as the area editor responsible for special issues of IEEE Signal Processing Magazine (2021-2023). She is a Fellow of IEEE, AAIA, and ELLIS.