

Robust Phase Linking in InSAR

Phan Viet Hoa Vu, Arnaud Breloy, Frédéric Brigui, Yajing Yan *Member IEEE*, and Guillaume Ginolhac *Senior Member, IEEE*

Abstract—Phase linking is a prominent methodology to estimate coherence and phase difference in interferometric synthetic aperture radar. This method is driven by a maximum likelihood estimation approach, which allows to fully exploit all the possible interferograms from a time series. Its performance is, however, known to be affected by the accuracy of the covariance matrix estimation step, which usually requires to introduce additional prior information on its structure when there is a small sample support (spatial window). Moreover, most phase linking algorithms are built upon the sample covariance matrix, due to the assumption of an underlying Gaussian distribution. In a scenario where SAR data is high resolution, or when the study area is spatially heterogeneous (e.g., urban area), this assumption can also limit the accuracy of the covariance matrix estimation step. Considering the two aforementioned issues, we introduce alternative statistical models, whose maximum likelihood estimators then yield new phase linking algorithms. In order to be robust to non-Gaussian data, we consider the use of a more general model of scaled mixture of Gaussian. To address small sample support issues, we also generalize this approach to a possibly low-rank structured covariance matrix. A unified algorithm to perform phase linking given these models is then derived and validated by simulations and a real data case (Sentinel-1 data).

Index Terms—interferometric synthetic aperture radar (InSAR), distributed scatterers (DS), phase linking (PL), maximum likelihood estimator (MLE), Scaled Gaussian distribution, covariance matrix, low-rank (LR).

I. INTRODUCTION

Multi-Temporal Interferometry SAR (MT-InSAR) analysis has become an useful tool to estimate deformation in sub-centimeters accuracy with low cost and over large coverage. The accuracy of the estimated deformation, however, is limited by target decorrelation. To overcome with this constraint, two groups of approaches were introduced: Permanent Scatterer Interferometry (PSI) and Distributed Scatterer Interferometry (DSI), corresponding to Permanent Scatterer (PS) and Distributed Scatterer (DS), respectively. PSI focuses on minimizing the signal decorrelation by using point-wise stable scatterers which are known as PS. Despite the preserved spatial resolution, this method is limited by its sparse PS points coverage particularly in natural areas. On the other hand, DSs are distributed over several resolution cells which endure the decorrelation over time. To reduce the target decorrelation, Small Baseline Subset (SBAS) approach in DSI uses small spatial and temporal baselines SAR image pairs. An example of this approach is given in [1]. Another important approach in DSI is Phase Linking (PL), or Phase Triangulation

Algorithm (PTA) [2]. The main idea of PL is to fully exploit all possible combinations of a SAR image stack, which is expected to improve Signal-to-Noise ratio (SNR) of single referenced phase estimation and thus, increase the accuracy of deformation retrieval. In this setup, the temporal and spatial decorrelation is generally properly tackled by weighting all interferograms in a maximum likelihood sense.

Indeed, phase linking was initially expressed as a phase estimation problem given an assumed proper (circular) multivariate model [2, 3]. Under this assumption, the phase differences (interferograms) directly appear in the structure of the covariance matrix of SAR images. The PL algorithm then aims to solve the maximum likelihood estimation of these phases for a given plug-in estimate of the coherence. Subsequent to PL, other frameworks integrated this maximum likelihood based scheme: an overview of the different objective functions and plug-in estimates is for example presented in [4]¹. Notably, low-rank structured covariance matrix estimates have been used to improve the estimation accuracy in CAESAR [5] and EMI [6]. This was motivated by the fact that DSI approaches degrade the spatial resolution because of the multi-looking applied during the interferogram computation. In principle, the number of samples within a multi-looking window should be at least twice the number of images in the time series dataset to ensure the accuracy of the covariance matrix estimation. Therefore, the trade-off between the quality of the covariance matrix and an improved spatial resolution is a problem for long time series. The low-rank structure of the covariance matrix is an assumption that is well-motivated from empirical measurements, and allows to improve this trade-off in practice. Alternatively, the sequential estimator [7] provides a way to process a long temporal SAR data stack by applying PL on small batches given by the time series data division.

In the aforementioned works, the phase estimation is related to an underlying Gaussian model. The plug-in estimates for PL are consequently built upon the sample covariance matrix. However, it is acknowledged that target decorrelation and atmospheric phase screen (APS) can lead to non-Gaussian statistics [8]. Moreover, this assumption can also be untrue in case of heterogeneous areas and high spatial resolution data. When dealing with heavy-tailed distributions, the sample covariance matrix is not an accurate estimate, which severely degrades the performance of PL-based schemes.

In this paper, we propose to tackle this issue by reformulating the phase estimation problem as a maximum likelihood of

P. V. H. Vu is with ONERA-DEMR, University Paris Saclay, and with LISTIC (EA3703), University Savoie Mont-Blanc. A. Breloy is with LEME (EA4416), University Paris Nanterre. F. Brigui is with ONERA-DEMR, University Paris Saclay. Y. Yan and G. Ginolhac are with LISTIC (EA3703), University Savoie Mont-Blanc.

¹Compared to SBAS approaches, PL-based methods must invert a full covariance matrix, which can be costly in terms of computational time. However, using the full covariance matrix provides sufficient redundancy and interferograms of various temporal baselines, which allows mitigating phase bias and increasing the accuracy of the phase estimates.

a larger class of statistical models. We consider modelling the samples as following a mixture of scaled Gaussian distributions. Indeed, such model is well known to provide a good empirical fit to high resolution radar data, and ensures robustness to a heavy-tailed distribution within the large family of complex elliptically symmetric distributions [9]. For this model, we additionally consider the possibility for the covariance matrix to be low-rank structured as in [4–6]. To perform phase linking under these assumptions, we consider the problem of joint maximum likelihood estimation of covariance coefficients and the phases differences (a slight reparameterization of the PL covariance structure, further discussed in the paper). This model was recently addressed in [10, 11] for the Gaussian case, and compared favorably to PL algorithms that rely on plug-in estimates. We therefore generalize this approach to mixture of scaled Gaussian distributions and structured covariance matrices. The optimization of the likelihood is tackled in an unified way using a block coordinate descent (BCD) and the majorization-minimization (MM) algorithm. We derive 4 algorithms that allow for investigating various options of SAR data model and structure of the covariance matrix: GPL solves the maximum likelihood for the Gaussian model and full-rank covariance matrix [10, 11], SGPL generalizes this approach to mixture of scaled Gaussian models, while GPL_{LR} and SGPL_{LR} correspond to their low-rank counterparts.

The proposed models and algorithms are validated by simulations and a real data case study. The real dataset is composed of a time series of Sentinel-1 images of Mexico City acquired from July 3rd, 2019 to December 18th, 2019. This area is of particular interest in InSAR, as the subsidence rate can reach up to 40 cm/year due to over exploitation of aquifers [12]. In both cases, our approaches are compared to the standard PL methodology.

The rest of the paper is organized as follows. Section II presents the mixture of scaled Gaussian model and the multi-temporal InSAR covariance matrix structure.

Section III discusses an alternate parameterization of the covariance structure, its practical interest, and its implication in terms of phase closure in InSAR. Section IV details the maximum likelihood algorithms derivations. Finally, Section V and VI display the synthetic data simulations and the real data experiments, respectively.

II. MODELING SAR IMAGE TIME SERIES

From a given stack of N co-registered SAR images, we consider a local patch of L multivariate pixels denoted as $\{\mathbf{x}_i\}_{i=1}^L$, with $\mathbf{x}_i \in \mathbb{C}^N$, $\forall i \in \llbracket 1, L \rrbracket$ (cf. Fig. 1). A multivariate pixel \mathbf{x}_i thus contains a local observation for N snapshots, i.e.

$$\mathbf{x}_i = [x_i^1, \dots, x_i^N]^T \in \mathbb{C}^N. \quad (1)$$

We will assume that the set $\{\mathbf{x}_i\}_{i=1}^L$ is a homogeneous patch containing L adjacent pixels with similar scattering and statistical properties. Thus $\{\mathbf{x}_i\}_{i=1}^L$ is a set of i.i.d. vectors that are realizations of the random variable \mathbf{x} . The aim of this section is to formalize several statistical models where phase differences appear as parameters. This will allow us to derive

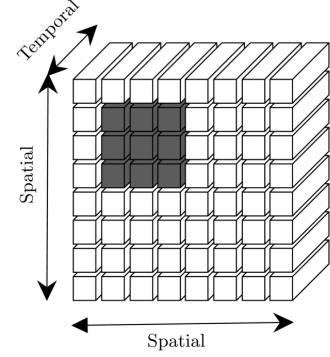


Fig. 1: Stack of N co-registered SAR images and multi-looking window: gray pixels represent the current local patch, denoted $\{\mathbf{x}_i\}_{i=1}^L$.

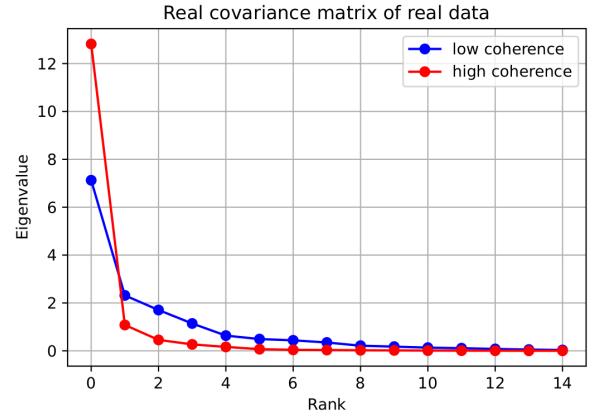


Fig. 2: Eigenvalues of the sample covariance matrix in low and high coherence areas of the considered Sentinel-1 dataset, with $N = 15$, $L = 64$.

a new phase difference estimation algorithm from a maximum likelihood approach in section IV.

A. Covariance structure

From the standard physical considerations of SAR interferometry, we assume the following first and second order moments relations:

$$\begin{aligned} \mathbb{E}[x^n] &= 0, \quad \forall n \in \llbracket 1, N \rrbracket \\ \mathbb{E}[x^k(x^l)^*] &= v_{k,l} \sigma_k \sigma_l e^{j(\theta_k - \theta_l)}, \quad \forall (k, l) \in \llbracket 1, N \rrbracket^2 \end{aligned} \quad (2)$$

where

- $\sigma_n^2 = \mathbb{E}[x^n(x^n)^H]$ is the variance of x^n . We denote the vector of standard deviations $\boldsymbol{\sigma} = [\sigma_1, \dots, \sigma_N]$.
- $v_{k,l} \in [0, 1]$ is the coherence coefficient between x^k and x^l . We denote $\boldsymbol{\Upsilon}$ the coherence matrix, with entries $[\boldsymbol{\Upsilon}]_{k,l} = v_{k,l}$. Also remark that $[\boldsymbol{\Upsilon}]_{l,l} = 1, \forall l \in \llbracket 1, N \rrbracket$.
- θ_n is the phase at instant n . We denote the phase vector $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]$, and the corresponding vector of complex arguments is

$$\mathbf{w}_{\boldsymbol{\theta}} = [e^{j\theta_1}, \dots, e^{j\theta_N}] \in \mathbb{T}_N, \quad (3)$$

where $\mathbb{T}_N = \{\mathbf{w} \in \mathbb{C}^N \mid |[\mathbf{w}]_i| = 1, \forall i \in \llbracket 1, N \rrbracket\}$ is the N -torus of phase-only complex vectors. By convention, we will use the reference $\theta_1 = 0$, which is equivalent to $[\mathbf{w}_\theta]_1 = 1$.

The covariance structure in (2) can then be rewritten in matrix form as

$$\mathbb{E} [\mathbf{x}\mathbf{x}^H] \stackrel{\Delta}{=} \mathbf{C} = \text{diag}(\mathbf{w}_\theta) \underbrace{((\sigma\sigma^\top) \odot \boldsymbol{\Upsilon})}_{\Psi} \text{diag}(\mathbf{w}_\theta)^H. \quad (4)$$

Where $\boldsymbol{\Psi}$ is coherence matrix scaled by the variance coefficients. We also can notice that this decomposition coincides with the modulus-argument decomposition, i.e.:

$$\mathbf{C} = \text{mod}(\mathbf{C}) \odot \arg(\mathbf{C}) \stackrel{\Delta}{=} \boldsymbol{\Psi} \odot (\mathbf{w}_\theta \mathbf{w}_\theta^H), \quad (5)$$

In the standard case, the matrix $\boldsymbol{\Psi}$ is assumed to have no specific structure besides being symmetric with positive entries. However, it is generally relevant to assume that it also exhibits a low-rank structure, i.e. that it can be decomposed as $\boldsymbol{\Psi} = \boldsymbol{\Psi}_R + \sigma^2 \mathbf{I}$, where $\boldsymbol{\Psi}_R$ is symmetric of rank R , and $\sigma^2 \in \mathbb{R}_+^*$ is the noise floor variance.

This assumption is related to low-rank linear models [13] and is often validated from empirical measurements. For example, Fig. 2 displays the spectrum of the sample covariance matrix for two areas of the Sentinel-1 dataset, where we can indeed observe a low-rank structure. Accounting for this structure in the phase estimation problem turns out to be beneficial in terms of phase estimation accuracy [4]. It also allows for improving the spatial resolution, i.e., estimating the phases with a reduced sample size L . In practice, the rank R can be tuned from observations, or estimated locally using model-order selection methods². Simulations in section V will also exhibit that the proposed methods are robust to slight mismatches of R in practice. Thus, we consider setting a single fixed rank R to process the whole data cube.

B. Scaled mixture Gaussian model for SAR images

Given the first and second order moment relations in (2), a common assumption is that $\{\mathbf{x}_i\}_{i=1}^L$ is distributed according to a complex circular (proper) Gaussian distribution, i.e., $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$. This yields the corresponding negative log-likelihood of the dataset

$$\mathcal{L}_G(\mathbf{C}) \propto \text{Tr} \{ \mathbf{C}^{-1} \mathbf{S} \} + \log |\mathbf{C}| + \text{const.} \quad (6)$$

with $\mathbf{S} = \frac{1}{L} \sum_{i=1}^L \mathbf{x}_i \mathbf{x}_i^H$. This model assumption is the base of the well known phase linking algorithm (detailed in the next section).

In this work, we propose to consider an alternative multivariate model. The motivation comes from empirical measurements exhibiting heavy tails. For example, Fig. 3 presents the histogram of the real part of a Sentinel-1 SLC inside a spatial window, where the empirical distribution is more in accordance with a Generalized Gaussian distribution. A practical and robust way to account for this type of measurements is to consider scaled Gaussian models. In this case,

²About rank estimation, the reader is referred to the overview [14], and more recent methods using shrinkage [15], random matrix theory [16], or M -estimators [17].

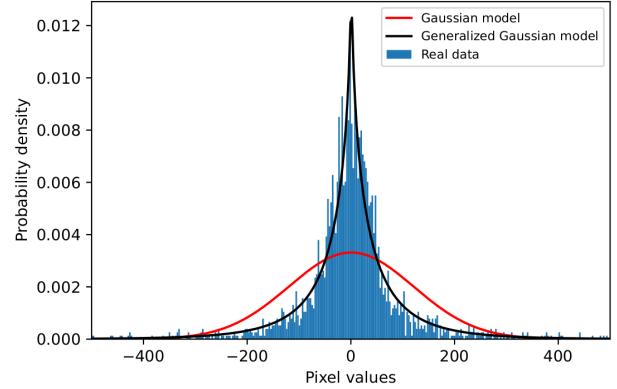


Fig. 3: Empirical distribution of the real part of Sentinel-1 SLC samples taken from a 50×50 window on the image acquired on Jul 3, 2019.

the data is modeled as Gaussian conditionally to an unknown deterministic scale, i.e. $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \tau_i \mathbf{C}), \forall i \in \llbracket 1, L \rrbracket$. This yields the corresponding negative log-likelihood of the dataset

$$\mathcal{L}_{SG}(\mathbf{C}) \propto \sum_{i=1}^L \left[\frac{\mathbf{x}_i^H \mathbf{C}^{-1} \mathbf{x}_i}{\tau_i} + \log |\tau_i \mathbf{C}| \right] + \text{const.} \quad (7)$$

A main interest of these models is that they allow for a robust fit (i.e., suited to any underlying distribution) over the whole family of complex elliptically symmetric distributions [9], that encompasses generalized Gaussian, multivariate- t , K -distribution, etc.

C. Phase linking algorithms and positioning of the contributions

Phase linking [2–4] refers to a maximum likelihood approach for the covariance structure (4)–(5) assuming the Gaussian model. More precisely, phase linking corresponds to an approximate maximum likelihood estimator of θ for a given prior estimate of the matrix $\boldsymbol{\Psi}$, denoted $\hat{\boldsymbol{\Psi}}$.

After several manipulations of the likelihood \mathcal{L}_G in (6) and the covariance structure in (4)–(5), the phase linking problem can be reformulated as

$$\begin{aligned} & \underset{\mathbf{w}_\theta}{\text{minimize}} \quad \mathbf{w}_\theta^H (\hat{\boldsymbol{\Psi}}^{-1} \odot \mathbf{S}) \mathbf{w}_\theta. \\ & \text{subject to} \quad \mathbf{w}_\theta \in \mathbb{T}_N \\ & \theta_1 = 0 \end{aligned} \quad (8)$$

that can be solved, e.g., with majorization-minimization fixed point iterations (cf. Section IV-C). At first, phase linking was proposed with $\hat{\boldsymbol{\Psi}} = \text{mod}(\mathbf{S})$, i.e., the modulus of the sample covariance matrix. Improvements were then brought by rather using plug-ins of low-rank approximations of this estimate (see, e.g., [4]).

Notice that, in this perspective, the phase-linking algorithm is bound to be a two-step approach that relies on an intermediate plug-in estimate of the variance and coherence coefficients (i.e., the matrix $\boldsymbol{\Psi}$). Most notably, there is, to the best of our knowledge, no explicit (nor tractable) joint

maximum likelihood estimators for the modulus-argument decomposition when assuming some additional phase and/or low-rank structure in the covariance matrix. This is because the modulus and argument are not holomorphic functions [18, 19], which probably explains why the current literature on phase linking is focused on two-steps algorithms (as also noticed in [10, 11]). To tackle this issue, the section III presents a suitable (i.e., differentiable for optimization) alternate decomposition of the problem, which was considered (though, not clearly explicated) within Gaussian models in [10, 11]. This parameterization allows deriving joint maximum likelihood estimation algorithms that generalize the phase linking approach to non-Gaussian distributions and structured covariance matrices, which is presented in section IV.

III. AN OPTIMIZATION-ORIENTED REPARAMETERIZATION OF THE COVARIANCE MATRIX

A. Real core and phase decomposition

We consider an alternate decomposition of the covariance matrix that relies on the following reparameterization of the second order moment:

$$\mathbb{E}[x^k(x^l)^*] = \gamma_{k,l} \sigma_k \sigma_l e^{j(\theta_k - \theta_l)}, \quad \forall (k, l) \in \llbracket 1, N \rrbracket^2, \quad (9)$$

where all the quantities from (2) hold, except for the coefficient $\gamma_{k,l}$, defined as:

- $\gamma_{k,l} \in [-1, 1]$ is the real correlation coefficient between x^k and x^l . We denote Γ the real correlation matrix, with entries $[\Gamma]_{k,l} = \gamma_{k,l}$, and where $[\Gamma]_{l,l} = 1, \forall l \in \llbracket 1, N \rrbracket$.

This yields, in matrix form, the following covariance structure

$$\mathbb{E}[\mathbf{x}\mathbf{x}^H] \triangleq \mathbf{C} = \text{diag}(\mathbf{w}_\theta) \underbrace{((\boldsymbol{\sigma}\boldsymbol{\sigma}^\top) \odot \Gamma)}_{\Sigma} \text{diag}(\mathbf{w}_\theta)^H. \quad (10)$$

where $\Sigma \in \mathcal{S}_N^{++}$ (\mathcal{S}_N^{++} is the space of $N \times N$ symmetric positive definite matrices). We refer to this parameter Σ as the *real core* of the covariance matrix. In order to account for a potential low-rank structure (cf. motivations in section II), we also consider the optional structure

$$\Sigma = \Sigma_R + \sigma^2 \mathbf{I}, \quad \text{with } \Sigma_R \in \mathcal{S}_{N,R}^+, \quad (11)$$

where $\mathcal{S}_{N,R}^+$ denotes the space $N \times N$ positive semi definite symmetric matrices of rank R .

To summarize, we denote the covariance matrix parameterization from (10) as

$$\mathbf{C}(\Sigma, \theta) \triangleq \text{diag}(\mathbf{w}_\theta) \Sigma \text{diag}(\mathbf{w}_\theta)^H = \Sigma \odot (\mathbf{w}_\theta \mathbf{w}_\theta)^H \quad (12)$$

where

- $\mathbf{w}_\theta \in \mathbb{T}_N$ is the vector of complex arguments in (3).
- Σ is the real core of the covariance matrix, which is either unstructured (in \mathcal{S}_N^{++}), or low-rank structured as in (11).

In section IV, we will rely on the core-correlation-phase decomposition from (10) (rather than the modulus-argument decomposition in (4)) to derive phase-difference estimation algorithms. The main underlying motivation is that \mathcal{S}_N^{++} (resp $\mathcal{S}_{N,R}^+$) and \mathbb{T}_N are smooth manifolds. This makes (10) more suited to the formalization and resolution of optimization problems, such as joint maximum likelihood estimation. Notice that

the two decompositions still coincide when $\Sigma_{i,j} \geq 0, \forall i, j \in \llbracket 1, N \rrbracket$ (i.e., $(\Sigma = \Upsilon)$). However, considering correlation rather than coherence coefficients (i.e., $\Sigma \in \mathcal{S}_N^{++}$) is more general and covers a larger matrix space. In the specific context of phase-difference estimation, it has an acknowledged cost of inducing potential ambiguities regarding the phase-closure property. This issue and several solutions are discussed in detail in section III-B. Finally, a clear practical interest of such approach is illustrated in the experiments of both synthetic simulations (Section V) and real data (Section VI).

B. Notes on phase closure ambiguities

From the model (2) we have the phase differences between two images indexed k and l defined as $\Delta_{k,l} = \theta_k - \theta_l$. Phase closure (or phase consistency) refers to the property

$$\Delta_{i,j} + \Delta_{j,k} + \Delta_{k,i} = 0 \quad (13)$$

that should be satisfied for all triplet $\{i, j, k\}$. Indeed, phase closure is an important property in MT-InSAR, as it is related to the obvious continuity of physical phenomena, such as Earth displacement. All DS based multi-temporal InSAR approaches are based on multi-looking (spatial average) to reduce the decorrelation noise, the phase closure is therefore not respected in most cases due to either statistical or physical properties of targets present in SAR images [20]. For example, given any estimate of the covariance matrix $\hat{\mathbf{C}}$, a direct approach to phase difference estimation is to produce estimates from (2) as

$$\hat{\Delta}_{k,l}(\hat{\mathbf{C}}) = \arg\{[\hat{\mathbf{C}}]_{k,l}\}, \quad (14)$$

which does not generally satisfy phase closure if such structure is not imposed in the estimation process of $\hat{\mathbf{C}}$.

Leveraging the phase closure (in particular that related to statistical properties of targets) within a redundant interferogram network has been the core of many multi-temporal InSAR approaches. For example, phase linking approaches produce estimates that satisfy naturally this property by the way of constructing the phase difference. SBAS approaches retrieve the phase closure based on a minimum variance estimator [1]. Recently, [21, 22] indicate that using only small baseline interferograms results in a phase bias related to physical properties of targets, whereas approaches that use the full covariance matrix of SAR images like phase linking can mitigate this bias. Several recent works like [23–25] propose some post-processing approaches to correct the small baseline interferogram induced phase bias and thus to retrieve the phase closure.

In the present work, a covariance matrix estimate will be obtained as $\mathbf{C}(\hat{\Sigma}, \hat{\theta})$ defined in (12), where the couple $\{\hat{\Sigma}, \hat{\theta}\}$ is an output of a maximum likelihood estimation algorithm (cf. Algorithm 1, derived in section IV).

A first option consists in computing phase difference estimates from (14) applied to the estimate $\mathbf{C}(\hat{\Sigma}, \hat{\theta})$. However, since

$$\gamma_{k,l} \sigma_k \sigma_l e^{j(\theta_k - \theta_l)} = (-\gamma_{k,l}) \sigma_k \sigma_l e^{j(\theta_k - \theta_l) + \pi} \quad (15)$$

provides two valid candidates for the decomposition (2) of the entry $[\mathbf{C}(\hat{\Sigma}, \hat{\theta})]_{k,l}$, this method can produce unexpected phase

	$\Sigma \in \mathcal{S}_N^{++}$	Σ LR as in (11)
Gaussian (\mathcal{L}_G)	GPL	GPL _{LR}
Scaled Gaussian (\mathcal{L}_{SG})	SGPL	SGPL _{LR}

TABLE I: Statistical models and corresponding acronyms

jumps³ of $\pm\pi$. We also notice that if some elements of the real-core estimate $\hat{\Sigma}$ are non-positive, the resulting estimates of phase differences do not satisfy the phase closure (conversely, if $\hat{\Sigma}$ has all positive entries, there is no ambiguity). The second option is thus to use the output complex argument estimate \mathbf{w}_{θ} in order to construct phase difference estimates as

$$\hat{\Delta}_{k,l}(\mathbf{w}_{\theta}) = \arg\{\mathbf{w}_{\theta}\mathbf{w}_{\theta}^H\}_{k,l}, \quad (16)$$

that satisfy the phase closure by construction.

When applied to real world data, we observed that the first option (i.e., (14) applied to the output estimate) provides cleaner interference maps.

Thus we advocate this simple solution when only focusing on phase difference between 2 given dates of the time series. For applications where the phase closure property is at stake, (16) offers a better alternative, but we found beneficial to perform a sanity check regarding the entries of the entries of $\hat{\Sigma}$. For example, when all entries are negative, it means that a constant π has been absorbed. In more complex cases (i.e., with ambiguities) an interesting solution comes from focusing on the first above-diagonal elements $[\mathbf{C}(\hat{\Sigma}, \hat{\theta})]_{k,k+1}$ and carefully integrating the phase shifts between each date.

IV. MAXIMUM LIKELIHOOD ESTIMATION IN SCALED GAUSSIAN MODELS

In this section, we propose an unified block-coordinate descent algorithm to solve for the maximum likelihood estimation problem under the various models defined in section II and III. The modeling options and corresponding acronyms are defined in Table I. First, remark that the likelihood \mathcal{L}_G is obtained as \mathcal{L}_{SG} with $\tau_i = 1, \forall i \in [1, L]$. Hence, we can focus on the generic problem

$$\begin{aligned} & \underset{\Sigma, \theta, \{\tau_i\}_{i=1}^L}{\text{minimize}} \quad \mathcal{L}_{SG}(\mathbf{C}(\Sigma, \theta)) \\ & \text{subject to} \quad \Sigma \in \mathcal{S}_N^{++} \\ & \quad \Sigma \text{ as in (11) (optional)} \\ & \quad \theta_1 = 0 \end{aligned} \quad (17)$$

and simply skip the block update of $\{\tau_i\}_{i=1}^L$ to solve for Gaussian maximum likelihood. The updates for the block-coordinate descent algorithm are detailed in the following subsections, and the corresponding algorithm is reported in the table Algorithm 1.

³We still note that this ambiguity appears to be inherent to the multi-temporal phase difference estimation problem, as we observed biases and phase jumps of $\pm\pi$ even with phase linking algorithms that rely only on plug-in matrices with only positive entries (see, e.g., phase histograms of PL algorithm in figure 4).

Algorithm 1 BCD algorithm for phase estimation

-
- 1: **Entry:** Samples $\{\mathbf{x}_i\}_{i=1}^L$, model choice in Table I
 - 2: Set $\tau_i = 1, \forall i \in [1, L]$ and $\tilde{\mathbf{S}} = \mathbf{S}$
 - 3: **repeat**
 - 4: **if** (model is \mathcal{L}_{SG})
 - 5: Update $\{\tau_i\}_{i=1}^L$ with (19)
 - 6: Update $\tilde{\mathbf{S}}$ with (21)
 - 7: Update Σ with (25)
 - 8: **if** (model is LR)
 - 9: Use projected update of Σ with (26)-(27)
 - 10: Call Algorithm 2 with $\mathbf{M} = \Sigma^{-1} \odot \tilde{\mathbf{S}}$
 - 11: Update θ from the output \mathbf{w} of Algorithm 2
 - 12: **until** Convergence
 - 13: **Output:** MLEs Σ, θ and $\mathbf{C}(\Sigma, \theta)$
-

A. Update $\{\tau_i\}_{i=1}^L$

For fixed Σ and θ , the matrix $\mathbf{C}(\Sigma, \theta)$ is constant and thus, will be denoted \mathbf{C} to lighten the exposition. Updating $\{\tau_i\}_{i=1}^L$ requires to solve the L separate sub-problems

$$\underset{\tau_i}{\text{minimize}} \quad \frac{\mathbf{x}_i^H \mathbf{C}^{-1} \mathbf{x}_i}{\tau_i} + N \log \tau_i. \quad (18)$$

This yields closed form solutions for the update as

$$\tau_i^* = \frac{\mathbf{x}_i^H \mathbf{C}^{-1} \mathbf{x}_i}{N}, \quad \forall i \in [1, L]. \quad (19)$$

For the next steps, the variables $\{\tau_i\}_{i=1}^L$ will remain constant. Hence it will be practical to introduce the notation

$$\mathcal{L}_{SG}(\mathbf{C}|\{\tau_i\}_{i=1}^L) = \text{Tr}\{\mathbf{C}^{-1} \tilde{\mathbf{S}}\} + \log |\mathbf{C}|, \quad (20)$$

with the conditional re-scaling of the samples

$$\tilde{\mathbf{S}} = \frac{1}{L} \sum_{i=1}^L \tilde{\mathbf{x}}_i \tilde{\mathbf{x}}_i^H, \quad \text{with} \quad \tilde{\mathbf{x}}_i = \mathbf{x}_i / \sqrt{\tau_i}. \quad (21)$$

B. Update Σ

First, note that we have the relations

$$\mathbf{C}^{-1}(\Sigma, \theta) = \text{diag}(\mathbf{w}_{\theta}) \Sigma^{-1} \text{diag}(\mathbf{w}_{\theta})^H \quad (22)$$

and

$$\log |\mathbf{C}(\Sigma, \theta)| = \log |\Sigma| \quad (23)$$

Thus, updating Σ for fixed θ and $\{\tau_i\}_{i=1}^L$ requires to solve the problem

$$\begin{aligned} & \underset{\Sigma}{\text{minimize}} \quad \text{Tr}\{\Sigma^{-1} \text{diag}(\mathbf{w}_{\theta})^H \tilde{\mathbf{S}} \text{diag}(\mathbf{w}_{\theta})\} + \log |\Sigma| \\ & \text{subject to} \quad \Sigma \in \mathcal{S}_N^{++} \\ & \quad \Sigma \text{ as in (11) (optional)} \end{aligned} \quad (24)$$

with $\tilde{\mathbf{S}}$ defined in (21). If no spectral structure is imposed on Σ , the minimizer is then obtained as the real part of the modified sample covariance matrix

$$\Sigma^* = \text{real}(\text{diag}(\mathbf{w}_{\theta})^H \tilde{\mathbf{S}} \text{diag}(\mathbf{w}_{\theta})). \quad (25)$$

Alternatively, if we impose the additional structure as in (11), the solution is given as [13]

$$\Sigma^* = \text{real}\left(\mathcal{P}_R\left\{(\text{diag}(\mathbf{w}_{\theta})^H \tilde{\mathbf{S}} \text{diag}(\mathbf{w}_{\theta}))\right\}\right) \quad (26)$$

where \mathcal{P}_R is the projection operator on the set of matrices structured as $\mathcal{S}_{N,R}^+$ plus scaled identity:

$$\begin{aligned}\mathcal{P}_R : \Sigma &\stackrel{\text{EVD}}{=} \mathbf{U} \text{diag}(\mathbf{d}) \mathbf{U}^H \mapsto \mathbf{U} \text{diag}(\bar{\mathbf{d}}) \mathbf{U}^H \\ \text{with } \bar{\mathbf{d}} &= [d_1, \dots, d_R, \bar{d}, \dots, \bar{d}] \\ \text{and } \bar{d} &= \sum_{i=R+1}^N d_i / (N - R)\end{aligned}\quad (27)$$

C. Update θ

From the relations (22), (23), and

$$\text{Tr} \left\{ \text{diag}(\mathbf{w}_\theta) \Sigma^{-1} \text{diag}(\mathbf{w}_\theta)^H \tilde{\mathbf{S}} \right\} = \mathbf{w}_\theta^H (\Sigma^{-1} \odot \tilde{\mathbf{S}}) \mathbf{w}_\theta, \quad (28)$$

the problem of updating θ with fixed Σ and $\{\tau_i\}_{i=1}^L$ reads

$$\begin{aligned}\underset{\theta}{\text{minimize}} \quad & \mathbf{w}_\theta^H (\Sigma^{-1} \odot \tilde{\mathbf{S}}) \mathbf{w}_\theta \\ \text{subject to} \quad & \mathbf{w}_\theta = [e^{j\theta_1}, \dots, e^{j\theta_N}] \\ & \theta_1 = 0\end{aligned}\quad (29)$$

which is a modified instance of the phase linking problem as stated in (8). For the sake of completeness, the remaining of this section briefly presents the derivations of [11] that proposed a majorization-minimization algorithm to solve this phase linking step.

For any phase linking problem, we need to solve the generic formulation

$$\underset{\mathbf{w} \in \mathbb{T}_N}{\text{minimize}} \quad \mathbf{w}^H \mathbf{M} \mathbf{w} \quad (30)$$

with $\mathbf{M} \succcurlyeq \mathbf{0}$. The resolution of (30) then yields a solution θ^* for (29) from the phases of a solution \mathbf{w}^* of (30), when plugging $\mathbf{M} = \Sigma^{-1} \odot \mathbf{S}$. Also remark that the objective function in (30) is invariant to a constant phase-shift of all entries, thus the constraint $\theta_1 = 0$ can be achieved by subtracting θ_1 to all the updated phases a posteriori. To solve the problem (30), we will use the majorization-minimization (MM) framework: the MM algorithm is an iterative optimization procedure that operates with two steps⁴: *i*) (*majorization*) at current point \mathbf{w}_t find a surrogate function $g(\cdot | \mathbf{w}_t)$ so that it is tangent to the objective, $f(\mathbf{w}_t) = g(\mathbf{w}_t | \mathbf{w}_t)$, and majorizes it, $f(\mathbf{w}) \leq g(\mathbf{w} | \mathbf{w}_t)$, $\forall \mathbf{w} \in \mathbb{T}_N$; *ii*) (*minimization*) obtaining the next iterate as $\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbb{T}_N}{\text{argmin}} g(\mathbf{w} | \mathbf{w}_t)$. This algorithm enjoys nice convergence properties [26], notably a monotonic decrement of the objective function at each step. More details on this framework can be found in [27, 28]. The main interest of this approach is that it can yield a sequence of sub-problems that are easily solved (e.g., in closed-form). The two MM steps are derived for problem (30) in the following.

Notice that if we restrict \mathbf{w} to the constrained set \mathbb{T}_N , we have the relation

$$\mathbf{w}^H (\mathbf{M} - \lambda_{\max}^{\mathbf{M}} \mathbf{I}) \mathbf{w} = \mathbf{w}^H \mathbf{M} \mathbf{w} - \underbrace{N \lambda_{\max}^{\mathbf{M}}}_{\text{const.}}, \quad \forall \mathbf{w} \in \mathbb{T}_N \quad (31)$$

so optimizing either the objective in (30) or (31) will lead to the same solution. The quadratic form $\mathbf{w}^H (\mathbf{M} - \lambda_{\max}^{\mathbf{M}} \mathbf{I}) \mathbf{w}$ is concave so that it can be majorized at point \mathbf{w}_t by its first order Taylor expansion

$$g(\mathbf{w} | \mathbf{w}_t) = 2\Re\{\mathbf{w}^H \underbrace{(\mathbf{M} - \lambda_{\max}^{\mathbf{M}} \mathbf{I}) \mathbf{w}_t}_{-\tilde{\mathbf{w}}_t}\} + \text{const.} \quad (32)$$

⁴Notations in this short introduction match problem (30) for convenience.

Algorithm 2 MM algorithm for Phase-linking problem (29)

- 1: **Entry:** $\mathbf{M} \in \mathbb{C}^{N \times N}$, $\mathbf{w}_1 \in \mathbb{T}_N$ (starting point)
 - 2: **repeat**
 - 3: Compute $\tilde{\mathbf{w}}_t = (\lambda_{\max}^{\mathbf{M}} \mathbf{I} - \mathbf{M}) \mathbf{w}_t$
 - 4: Update $\mathbf{w}_t = \mathcal{P}_{\mathbb{T}_N}\{\tilde{\mathbf{w}}_t\}$
 - 5: $t = t + 1$
 - 6: **until** convergence
 - 7: **Output:** $\mathbf{w} \in \mathbb{T}_N$
-

minimizing this surrogate corresponds to the problem

$$\underset{\mathbf{w} \in \mathbb{T}_N}{\text{maximize}} \quad 2\Re\{\mathbf{w}^H \tilde{\mathbf{w}}_t\} \quad (33)$$

whose solution is $\mathbf{w}_{t+1} = \mathcal{P}_{\mathbb{T}_N}\{\tilde{\mathbf{w}}_t\}$, where $\mathcal{P}_{\mathbb{T}_N}\{\cdot\}$ is the operator that projects each entry of a vector on the unit sphere (i.e., entry-wise normalization). Hence, we have a majorization-minimization algorithm to solve for (30). The algorithm is summed up in the table Algorithm 2.

V. SIMULATIONS

A. Simulation parameters

We simulate two scenarios: a short and a longer temporal dataset of respectively $N = 5$ and $N = 15$ images. The real core Σ is set as a Toeplitz matrix representing the temporal decorrelation: we use a coherence of $\rho = 0.7$, i.e. $[\Sigma]_{i,j} = \rho^{|i-j|}$ (so $\gamma_{i,j} = \rho^{|i-j|}$ and $\sigma_i = 1$). Phase differences are generated linearly from 0 to 2 rad, i.e. $\Delta_{i+1,i} = \theta_{i+1} - \theta_i = 2/N$ rad. The resulting covariance matrix \mathbf{C} is obtained with (12). The L i.i.d. samples are simulated as $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \tau_i \mathbf{C})$. In the Gaussian case, $\tau_i = 1$, $\forall i \in \llbracket 1, L \rrbracket$. In the non-Gaussian case, we sample each τ_i according to a Gamma distribution $\tau \sim \Gamma(\nu, 1/\nu)$, which yields a multivariate K -distribution [9]. In this setup, $\nu \in \mathbb{R}^+$ pilots the tails of the distribution: $\nu \rightarrow \infty$ yields the Gaussian distribution as a limit case, while $\nu \rightarrow 0$ makes the distribution more heavy-tailed.

We assess the performance of the 4 proposed approaches listed in Table I. The standard PL and 2-pass InSAR (2-p InSAR) are used as baselines for comparison. It is recalled that PL is the usual 2-step plug-in MLE approximation, where $\hat{\Sigma} = \text{mod}(\mathbf{S})$ is used as plug-in of Σ to solve for the PL problem (8). This problem is here solved using the MM procedure in Algorithm 2. The 2-p InSAR is a direct (unstructured) estimate, where the phase differences are estimated from the averaged pixel values given the spatial window.

As criterion, we consider the mean squared error (MSE) on the estimation of entries of θ , which is further investigated by displaying histograms of estimated phase θ at each date. All mean squared errors and histograms are computed using 1000 Monte-Carlo trials.

B. Simulation results

We first analyse the set-up with a short time series of $N = 5$ dates. Fig. 4 presents the histograms of phase estimates for a sample size of $L = N + 1$, and the corresponding MSEs of the different estimates for the phase vector element θ_2 when L increases. This display is analysed for 3 setups (from

left to right): the Gaussian case, the K -distribution case with respectively $\nu = 1$ (heavy-tailed case), and $\nu = 0.1$ (extremely heavy-tailed). Notice that in this low-dimension setting, low-rank methods are not particularly beneficial, and thus, not displayed. Upon considering histograms of phase difference estimates obtained with a small temporal baseline (Fig. 4), both SGPL and 2-p InSAR exhibit superior performance compared to the other methods, irrespective of the underlying sample distribution. On the other hand, the estimations produced by PL and GPL tend to deviate from the true phase value when the samples exhibit a heavier-tailed distribution. With an underlying Gaussian distribution, the phase difference estimations from all approaches yield similar accuracy. We notice that 2-p InSAR, despite of its simplicity, provides good results compared to other multi-temporal InSAR approaches in the low sample support. This is due to the small time series size that limits the benefit of multi-temporal approaches. In case of non-Gaussian models ($\nu = 1$ and $\nu = 0.1$), we observe the robustness of SGPL to underlying heavy-tailed distributions: the heavier the tail of the K -distribution, the better the performance we can achieve with SGPL compared to the others.

For a larger time-series ($N = 15$), we investigate the impact of the temporal decorrelation by analysing the MSEs on $\{\theta_2, \theta_4, \theta_7\}$ with respect to increasing L . Fig. 5 displays the results for an underlying Gaussian distribution, while Fig. 6 displays the same results with an underlying heavy-tail distribution (K -distribution with $\nu = 1$). In the case of Gaussian samples, differences in terms of MSE are mainly observed with low sample support ($L < 40$). When L is larger than 40, multi-temporal InSAR approaches yield similar performance. The performance of 2-p InSAR degrades quickly when the temporal decorrelation increases. The performance of GPL and that of SGPL are comparable and they outperform the state of the art. In the case of non-Gaussian samples (Fig. 6), SGPL always provides the best results in terms of MSE and appears more robust to a lower temporal coherence. Furthermore, the use of LR structured covariance matrix in both cases introduces a significant improvement at low sample support, which implies the possibility of reducing the multi-looking window size, hence preserving more spatial resolution, especially in the case of long time series data. Note in particular that, in the simulations, the true covariance matrix does not have an exact LR structure; but GPL_{LR} and SGPL_{LR} still provide significant improvement at low sample support compared to GPL and SGPL.

Finally, Fig. 7 assesses the robustness of the low-rank methods with respect to the choice of the rank R . As in the previous simulation, the true covariance matrix of the tested scenario is not low-rank. However, we can still observe that the dimension reduction approach is beneficial in terms of MSE at low sample support L . At larger L , the performance of the low-rank estimate also remains close to the actual maximum likelihood estimator (full-rank SGPL) for a reasonable range of R , which illustrates that the method is also not critically sensitive to a rank mismatch.

VI. REAL DATA

A. Study area and dataset

Located at the South of Mexico Valley and surrounded by mountains, Mexico City is the largest and most populous city of Mexico. The rapid urbanisation raises ground water demand which makes extraction rate is twice higher than the recharge rate [29]. Around 70% of the water supply of Mexico City comes from aquifers [30]. Ground water over-exploitation generates compression of the aquifer system, leading to land subsidence [31, 32]. The maximum subsidence velocity reached up to 40 cm/yr [12]. Mexico City subsidence provides thus an interesting testing ground for MT-InSAR approaches.

We investigate two stacks of 5 and 15 images corresponding to a 2-month and a 6-month Sentinel-1 SLC datasets, which are acquired in descending orbit over the period from July 3rd, 2019 to August 20th, 2019 and to December 18th, 2019, respectively. These images are 12 days apart and cover approximately $14 \times 22 \text{ km}^2$ around the Mexico City.

Before applying any MLE-PL approaches, we co-registered our dataset with reference to the first date (Jul 3, 2019), then, topographic and orbital fringes are removed using SNAP software⁵. The next step is to use our proposed approaches to estimate wrapped phase differences (referred to the first date) from the full covariance of SAR image time series. For this experiment, we will focus on the improvement of the estimation of wrapped phase differences compared to the state-of-the-art PL approaches (i.e., the problem of retrieving the subsidence rate of the Mexico City from those phase-differences is out of the scope of this paper).

B. Real data results

In both scenarios of short and longer time series, only the longest temporal baseline interferograms are shown here with the objective to highlight the performance of the proposed approaches against the temporal decorrelation; the latter is one of the most important factors that degrade the quality of the phase difference estimation. Multi-looking window size is the same in all cases with $L = 8 \times 8$ pixels. For LR structures, the rank R is fixed to be 1 for all algorithms. The choice of R is based on the analysis of eigenvalues of the sample covariance matrix of the whole dataset (Fig. 2). According to Fig. 2, the first eigenvalue is sufficiently informative. In addition, a test of different rank values has also been performed to confirm the choice of $R = 1$.

Fig. 8 shows the longest temporal baseline interferogram (Jul 3, 2019 - Aug 20, 2019) estimated by PL, GPL and SGPL in case of $N = 5$ and $L = 64$ pixels with a full-rank and a low-rank model. A slight improvement in terms of noise reduction obtained with SGPL_{LR} compared to PL_{LR} can be observed in a zoom area on the top line of Fig. 10.

Similar to the synthetic simulations, since the time series is small (longest time span is 48 days), the impact of the temporal decorrelation is small, the performances of PL, GPL and SGPL are similar.

⁵<http://step.esa.int>

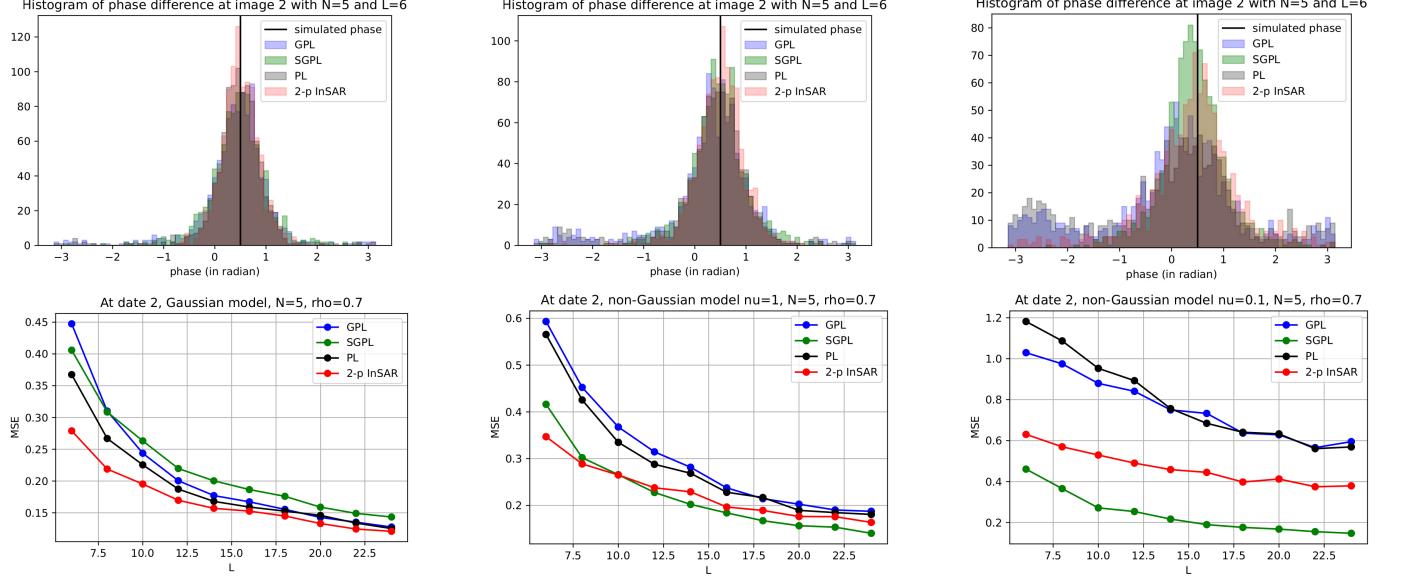


Fig. 4: Top row: histograms of phase estimators at date 2 for $N = 5$, $L = 6$. Bottom row: corresponding evolution of the MSEs on the estimation of the phase vector element θ_2 w.r.t. increasing L . From left to right column, the underlying distribution goes from Gaussian to heavy-tailed (i.e., K -distribution with $\nu = 1$ in the center, and $\nu = 0.1$ on the right).

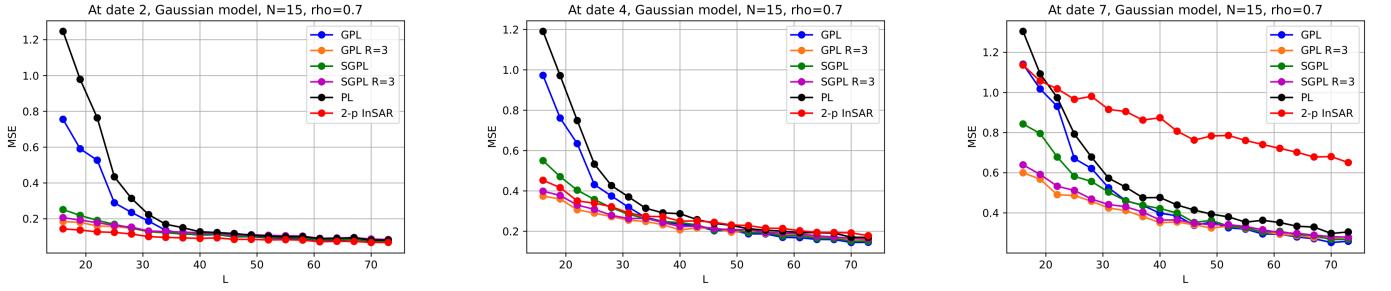


Fig. 5: Corresponding MSEs on $\{\theta_2, \theta_4, \theta_7\}$ with respect to increasing L (also from left to right column respectively). These figures display the results for an underlying Gaussian model, while the heavy-tailed case is in Fig. 6.

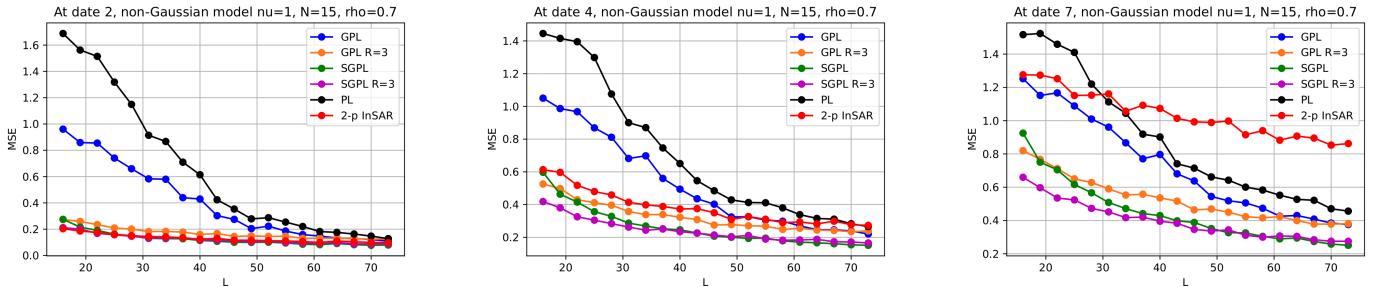


Fig. 6: Corresponding MSEs on $\{\theta_2, \theta_4, \theta_7\}$ with respect to increasing L (also from left to right column respectively). These figures display the results for an underlying K -distribution with $\nu = 1$, while the Gaussian case is in Fig. 5.

In case of longer time series (Fig. 9), the improvement in terms of noise reduction is more pronounced with the scaled Gaussian model and the LR structure. The benefit of scaled Gaussian model is particularly highlighted when we compare the interferogram obtained with PL_{LR} (right top) to that obtained with $SGPL_{LR}$ (right bottom). Indeed, with a longer temporal dataset, it is crucial to use the LR model to reduce the multi-looking window size (remind that the multi-

looking window size should be increased with the time series size in order to guarantee the performance of the covariance matrix estimation). To further showcase the performance of the SGPL approach, Fig. 10 shows a zoom of a focused area for a comparison between PL_{LR} and $SGPL_{LR}$ in cases of $N = 5$ and $N = 15$. A significant noise reduction is obtained with SGPL, especially with $N = 15$, which confirms the interest of the SGPL approach.

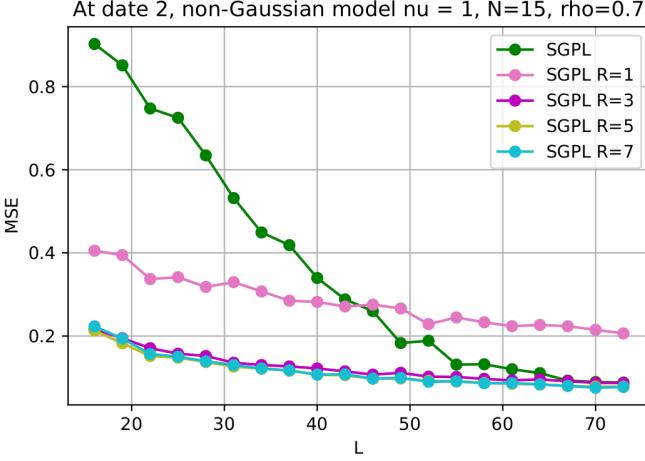


Fig. 7: MSEs on θ_2 with respect to L for SGPL_{LR} using different ranks R . K -distribution with $\nu = 1$, $N = 15$.

VII. CONCLUSIONS

In this work, novel robust-phase-linking algorithms are proposed to estimate wrapped interferometric phases with respect to a common date. Generalizing from the standard MLE-PL approaches, the proposed method relies on the consideration of scaled Gaussian statistics and the utilization of a LR structure of the covariance matrix. From this model, we derived an algorithm to compute the corresponding maximum likelihood estimator, which does not rely on plug-in estimates of the coherence. Simulations and applications to real data show the interest of considering this model: scaled Gaussian models allow for a significant gain in terms of noise reduction on wrapped interferograms, while the LR structure allows for reducing the number of needed samples with respect to the length of the image stack. Future works will investigate a streaming implementation of the proposed method, which will allow reducing the computational cost when processing large time series and integrating timely new images that arrive over time.

REFERENCES

- [1] Antonio Pepe, Yang Yang, Mariarosaria Manzo, and Riccardo Lanari, “Improved EMCF-SBAS processing chain based on advanced techniques for the noise-filtering and selection of small baseline multi-look DInSAR interferograms,” *IEEE transactions on geoscience and remote sensing*, vol. 53, no. 8, pp. 4394–4417, 2015.
- [2] Andrea Monti Guarneri and Stefano Tebaldini, “On the Exploitation of Target Statistics for SAR Interferometry Applications,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3436–3443, 2008.
- [3] Alessandro Ferretti, Alfio Fumagalli, Fabrizio Novali, Claudio Prati, Fabio Rocca, and Alessio Rucci, “A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 9, pp. 3460–3470, 2011.
- [4] Ning Cao, Hyongki Lee, and Hahn Chul Jung, “Mathematical Framework for Phase-Triangulation Algorithms in Distributed-Scatterer Interferometry,” *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1838–1842, 2015.
- [5] Gianfranco Fornaro, Simona Verde, Diego Reale, and Antonio Pauciullo, “CAESAR: An Approach Based on Covariance Matrix Decomposition to Improve Multibaseline–Multitemporal Interferometric SAR Processing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2050–2065, 2015.
- [6] Homa Ansari, Francesco De Zan, and Richard Bamler, “Efficient Phase Estimation for Interferogram Stacks,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 7, pp. 4109–4125, 2018.
- [7] Homa Ansari, Francesco De Zan, and Richard Bamler, “Sequential Estimator: Toward Efficient InSAR Time Series Analysis,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5637–5652, 2017.
- [8] Andrea Monti Guarneri and Stefano Tebaldini, “Hybrid Cramér–Rao Bounds for Crustal Displacement Field Estimators in SAR Interferometry,” *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 1012–1015, 2007.
- [9] Esa Ollila, David E Tyler, Visa Koivunen, and H Vincent Poor, “Complex elliptically symmetric distributions: Survey, new results and applications,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 5597–5625, 2012.
- [10] Chisheng Wang, Xiang-Sheng Wang, Yaping Xu, Bochen Zhang, Mi Jiang, Siting Xiong, Qin Zhang, Weidong Li, and Qingquan Li, “A New Likelihood Function for Consistent Phase Series Estimation in Distributed Scatterer Interferometry,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.
- [11] Phan Viet Hoa Vu, Frédéric Briguji, Arnaud Breloy, Yajing Yan, and Guillaume Ginolhac, “A New Phase Linking Algorithm for Multi-temporal InSAR based on the Maximum Likelihood Estimator,” in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022, pp. 76–79.
- [12] Yajing Yan, Marie-Pierre Doin, Penélope Lopez-Quiroz, Florence Tupin, Bénédicte Fruneau, Virginie Pinel, and Emmanuel Trouve, “Mexico city subsidence measured by insar time series: Joint analysis using ps and sbas approaches,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1312–1326, 2012.
- [13] Michael E Tipping and Christopher M Bishop, “Probabilistic principal component analysis,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.
- [14] Petre Stoica and Yngve Selén, “Model-order selection: a review of information criterion rules,” *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.
- [15] Lei Huang and Hing-Cheung So, “Source Enumeration Via MDL Criterion Based on Linear Shrinkage Estimation of Noise Subspace Covariance Matrix,” *IEEE Transactions Signal Processing*, vol. 61, no. 19, pp. 4806–4821, 2013.
- [16] Eugénie Terreaux, Jean-Philippe Ovarlez, and Frédéric Pascal, “Robust model order selection in large dimensional elliptically symmetric noise,” *arXiv preprint arXiv:1710.06735*, 2017.
- [17] Gordana Drašković, Arnaud Breloy, and Frédéric Pascal, “On the asymptotics of maronna’s robust PCA,” *IEEE Transactions on Signal Processing*, vol. 67, no. 19, pp. 4964–4975, 2019.
- [18] Walter Rudin, *Real and Complex Analysis*, 3rd Ed., McGraw-Hill, Inc., 1987.
- [19] Walter Appel and Emmanuel Kowalski, *Mathematics for physicists*, vol. 47, Princeton University Press Princeton, NJ, 2007.
- [20] Simon Zwieback, Xingyu Liu, Sofia Antonova, Birgit Heim, Annett Bartsch, Julia Boike, and Irena Hajnsek, “A statistical test of phase closure to detect influences on DInSAR deformation estimates besides displacements and decorrelation noise: Two case studies in high-latitude regions,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 9, pp. 5588–5601, 2016.
- [21] Homa Ansari, Francesco De Zan, and Alessandro Parizzi, “Study of systematic bias in measuring surface deformation with SAR interferometry,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 2, pp. 1285–1301, 2020.
- [22] Homa Ansari, Francesco De Zan, and Alessandro Parizzi, “Fading signal: An overlooked error source for distributed scatterer interferometry,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 3181–3184.
- [23] Francesco Falabella and Antonio Pepe, “On the phase nonclosure of multilook SAR interferogram triplets,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.
- [24] Yujie Zheng, Heresh Fattah, Piyush Agram, Mark Simons, and Paul Rosen, “On closure phase and systematic bias in multilooked SAR interferometry,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2022.
- [25] Yasser Maghsoudi, Andrew J Hooper, Tim J Wright, Milan Lazecky, and Homa Ansari, “Characterizing and correcting phase biases in short-term, multilooked interferograms,” *Remote Sensing of Environment*, vol. 275, pp. 113022, 2022.
- [26] Meisam Razaviyayn, Mingyi Hong, and Zhi-Quan Luo, “A Unified Convergence Analysis of Block Successive Minimization Methods for

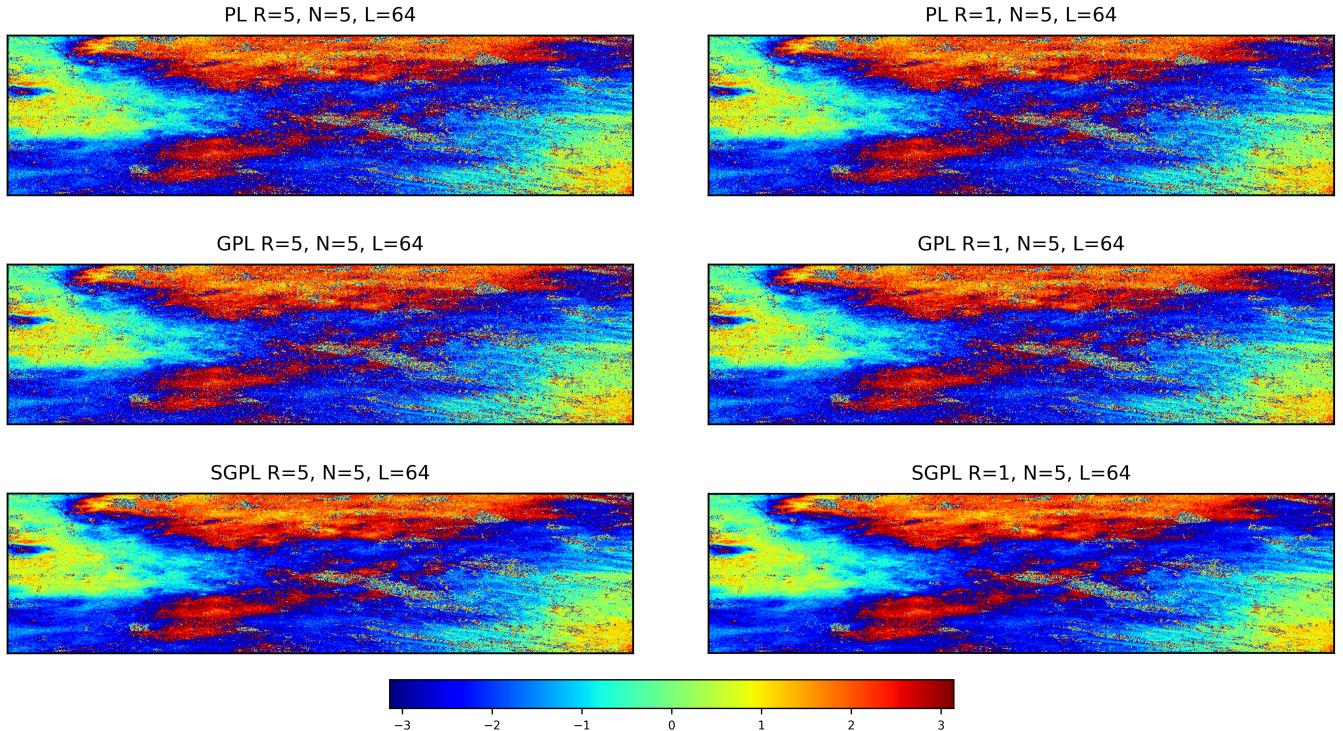


Fig. 8: The longest temporal baseline interferogram (Jul 3, 2019 - Aug 20, 2019) estimated by PL, GPL and SGPL (from top to bottom) in case of $N = 5$ and $L = 64$ pixels with a full-rank model (left) and a low-rank model (right)

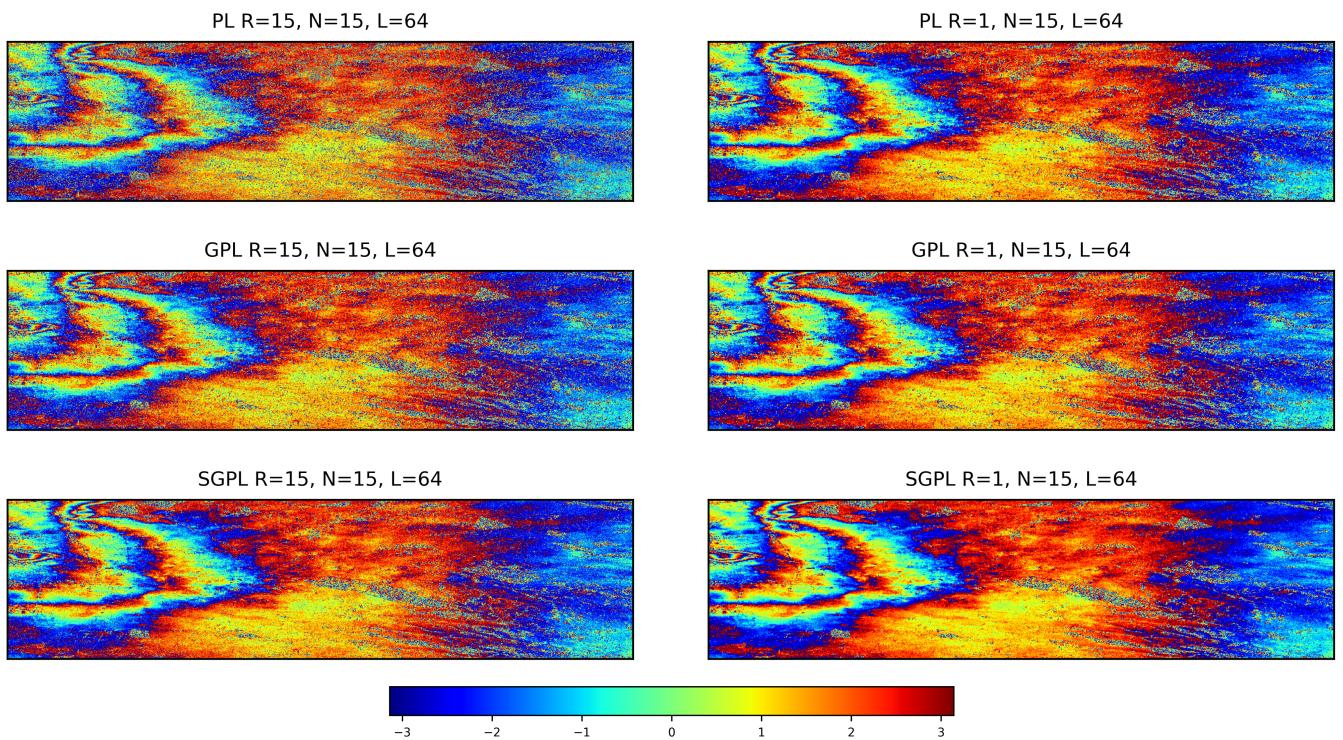


Fig. 9: The longest temporal baseline interferogram (Jul 3, 2019 - Dec 18, 2019) estimated by PL, GPL and SGPL (from top to bottom) in case of $N = 15$ and $L = 64$ pixels with a full-rank model (left) and a low-rank model (right)

Nonsmooth Optimization,” *SIAM on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.
[27] David R Hunter and Kenneth Lange, “A tutorial on MM algorithms,”

The American Statistician, vol. 58, no. 1, pp. 30–37, 2004.
[28] Ying Sun, Prabhu Babu, and Daniel P. Palomar, “Majorization-minimization algorithms in signal processing, communications, and

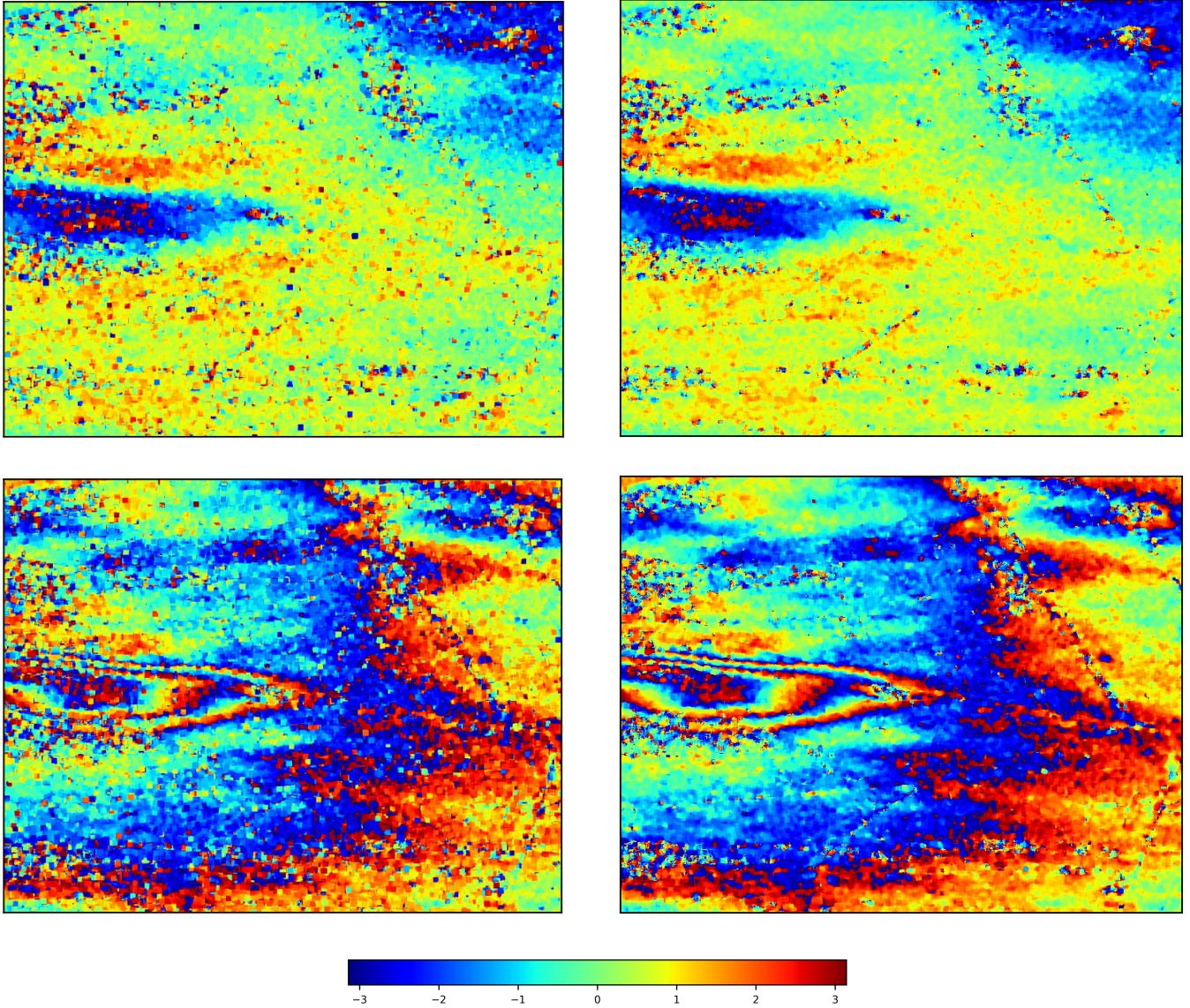


Fig. 10: A focused view for a comparison between PL_{LR} (left column) and SGP_{LR} (right column) in both cases $N = 5$ (top row) and $N = 15$ (bottom row).

- machine learning,” *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2016.
- [29] Cecilia Tortajada, “Challenges and realities of water management of megacities: The case of Mexico City Metropolitan Area,” *Journal of International Affairs*, vol. 61, no. 2, pp. 147–166, 2008.
 - [30] Cecilia Tortajada, “Water Management in Mexico City Metropolitan Area,” *International Journal of Water Resources Development*, vol. 22, no. 2, pp. 353–376, 2006.
 - [31] Penélope López-Quiroz, Marie-Pierre Doin, Florence Tupin, Pierre Briole, and Jean-Marie Nicolas, “Time series analysis of Mexico City subsidence constrained by radar interferometry,” *Journal of Applied Geophysics*, vol. 69, no. 1, pp. 1–15, 2009, Advances in SAR Interferometry from the 2007 Fringe Workshop.
 - [32] Estelle Chaussard, Shimon Wdowinski, Enrique Cabral-Cano, and Falk Amelung, “Land subsidence in central Mexico detected by ALOS InSAR time-series,” *Remote Sensing of Environment*, vol. 140, pp. 94–106, 2014.