Interferometric phase linking with Riemannian distances on covariance matrices

Elena Grosso, ONERA & CNAM
Florent Bouchard, CNRS, L2S
Frederic Brigui, ONERA
Guillaume Ginolhac, USMB
Arnaud Breloy, CNAM

***Abstract*—Multi-temporal interferometric SAR allows one to monitor Earth surface displacement from SAR image time series. In this context, Interferometric Phase Linking (IPL) is a technique used to denoise the phase of SAR images by leveraging all possible pairs of interferogram within the time series. The task can be reformulated as a covariance matrix fitting problem, where the aim is to recover the expected InSAR phase structure from a noisy estimate of the covariance matrix of a pixel patch. Such framework leaves open a choice regarding the matrix distance that will define the notion of "optimal fitting". Existing methods from the state of the art have mostly focused on maximum-likelihood and least-squares fitting formulations. In this paper, we explore the use of several Riemannian distances on the space of covariance matrices (affine invariant, log-euclidean, and Bures-Wasserstein) for IPL. We derive an optimization algorithm to solve the corresponding fitting problems, and simulations illustrate the interest of these distances in terms of estimation accuracy and computational complexity.**

***Index Terms*—Interferometric SAR, interferometric phase linking, covariance matrix estimation, Riemannian geometry.**

I. Introduction

In recent years, Synthetic Aperture Radar (SAR) missions have provided data consistently and systematically throughout long periods of time. The availability of SAR image time-series has lead to the introduction of Multi-Temporal Interferometric SAR (MT-InSAR) techniques, which allowed for high-precision estimation of terrain displacement throughout time (in the order of sub-centimeters). Among these techniques, Interferometric Phase-Linking (IPL) or Phase Triangulation has become a staple method to process zones with distributed scatterers.

The fundamental idea behind IPL techniques is to leverage the redundancy of the time series in order to compensate for the coherence loss witnessed between the images over time. Hence, these methods produce a (local) estimate of the phase of each image by leveraging the complex correlation coefficients (which gather both the coherence and phase difference information) between all possible pairs of images within the time series (cf. overview in [1]). Seminal works that introduced the IPL methodology [2], [3] were based on performing a maximum likelihood estimation (MLE) while assuming a complex circular Gaussian model and a given (plug-in) estimate of the coherence matrix. Subsequent developments were made by refining the plug-in estimate of the coherence matrix [4], [5], deriving algorithms robust to non-Gaussian distributions [6]–[8], or deriving numerically efficient approximate solutions [9].

The IPL problem was also reframed as a covariance fitting problem (COFI-PL) in [10]. This formulation corresponds to the previous MLE approaches when the KL divergence is chosen as the matrix distance. However, several works have also shown that considering other distances can be beneficial for IPL [10]–[12]. In this line of work, this paper proposes to leverage Riemannian distances on the space of covariance matrices [13], [14] for IPL. The contributions presented in this work are the following:

- We extend the COFI-PL framework of [10] to include several Riemannian distances on the space of covariance matrices, namely: the affine invariant, the log-euclidean, and the Bures-Wasserstein distances.
- We derive Riemannian optimization algorithms on the torus of phase-only complex vectors to solve the corresponding COFI-PL problems.
- Experiments on simulated data illustrate the interest of the approach in terms of estimation accuracy and computational complexity.

The rest of the paper is organized as follows: Section II introduces the InSAR context and the corresponding covariance matrix structure. Section III presents the COFI-PL framework, and the proposed new distances to perform IPL. Section IV presents the derivation of the optimization algorithms that solve the corresponding problems. Finally, section V presents the simulation study that validates the proposed methods.

II. InSAR covariance matrix structure

From a given datacube of p co-registered SAR images, we consider a local multivariate pixel patch $\{\mathbf{x}_i\}_{i=1}^n$, with $\mathbf{x}_i \in \mathbb{C}^p,\ \forall i \in [\![1, n]\!]$. Each sample $\mathbf{x}_i$ contains the complex-valued time-series (in chronological order) of one pixel over the p snapshots, i.e.

$$\mathbf{x}_i = \left[\ x_i^1,\ \cdots,\ x_i^p\ \right]^\top \in \mathbb{C}^p. \tag{1}$$

We assume that the patch is homogeneous, i.e., that the set $\{\mathbf{x}_i\}_{i=1}^n$ contains n pixels with similar scattering and statistical properties. From the standard physical considerations about distributed scatterers in InSAR, we can assume that the second order moment follows:

$$\arg\left(\mathbb{E}\left[x^q (x^\ell)^*\right]\right) = e^{j(\theta_q - \theta_\ell)},\ \forall (q,\ell) \in [\![1, p]\!]^2 \tag{2}$$

This work has received support from France 2030 through the project Académie Spatiale d'Île-de-France (https://academiespatiale.fr/) managed by the National Research Agency under bearing the reference ANR-23-CMAS-0041.

The relation (2) imposes that the modulus-argument decomposition of the covariance matrix satisfies

$$\mathbf{\Sigma} = \mathrm{mod}(\mathbf{\Sigma}) \circ \phi_{\mathbb{T}}(\mathbf{\Sigma}) \triangleq \mathbf{\Psi} \circ (\mathbf{w}\mathbf{w}^{\mathsf{H}}) \tag{3}$$

where $\phi_{\mathbb{T}} : x{=}re^{i\theta} \mapsto e^{i\theta}$ is the complex phase extraction operator (which extends to matrices by being applied entry-wise), and where $\mathbf{w}$ is a vector of p complex phases, i.e.:

$$\mathbf{w} \in \mathbb{T}_p = \{\mathbf{w} \in \mathbb{C}^p \mid |[\mathbf{w}]_q| = 1,\ \forall q \in [\![1, p]\!]\}. \tag{4}$$

The aforementioned covariance matrix structure is said to respect the *phase closure* property, since for all triplets $(q, \ell, j) \in [\![1, p]\!]^3$, we have

$$\arg(\mathbf{\Sigma}_{q\ell}) + \arg(\mathbf{\Sigma}_{\ell j}) + \arg(\mathbf{\Sigma}_{jq}) = 0. \tag{5}$$

This property is at the core of many multi-temporal InSAR algorithm developments.

III. Interferometric phase linking as a covariance fitting problem

A. Framework

Interferometric phase linking (IPL) consists in estimating the complex phase vector $\mathbf{w}$ from the sample set $\{\mathbf{x}_i\}_{i=1}^n$ while leveraging the phase closure property stated in (5) [1]. The class of covariance fitting phase linking (COFI-PL) [10] perform this estimation by fitting the structure (3) to any given plug-in estimate of the covariance matrix $\hat{\mathbf{\Sigma}}$. The problem is formulated as follows:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & d^2(\hat{\mathbf{\Sigma}}, \hat{\mathbf{\Psi}} \circ \mathbf{w}\mathbf{w}^{\mathsf{H}}) \\ \text{subject to} & \mathbf{w} \in \mathbb{T}_p \end{array} \tag{6}$$

where $\hat{\mathbf{\Sigma}}$ is a covariance matrix estimate, $\hat{\mathbf{\Psi}} \triangleq \mathrm{mod}(\hat{\mathbf{\Sigma}})$ denotes its modulus, and d is a matrix distance (or divergence). Note that, since the only variable is the vector $\mathbf{w}$, we will use the compact notation

$$f^d_{\hat{\mathbf{\Sigma}}}(\mathbf{w}) = d^2(\hat{\mathbf{\Sigma}}, \hat{\mathbf{\Psi}} \circ \mathbf{w}\mathbf{w}^{\mathsf{H}}) \tag{7}$$

for the objective in (6). The generic COFI-PL problem is thus finally expressed as

$$\underset{\mathbf{w} \in \mathbb{T}_p}{\text{minimize}} \quad f^d_{\hat{\mathbf{\Sigma}}}(\mathbf{w}). \tag{8}$$

It follows that many IPL algorithms can be proposed by setting $\hat{\mathbf{\Sigma}}$ and d. In this work, we will focus only on the choice of the matrix distance d. Hence, we will fix a common estimate $\hat{\mathbf{\Sigma}}$ to all methods. This estimator will be the regularized sample covariance matrix (RSCM), as it is one of the most common baseline of the IPL literature (cf. e.g. [5]), that is defined as

$$\hat{\mathbf{\Sigma}}(\beta) = \beta\mathbf{S} + (1-\beta)\frac{\mathrm{tr}(\mathbf{S})}{p}\mathbf{I}_p, \tag{9}$$

in which $\mathbf{S} = \frac{1}{n}\sum_{i=1}^N \mathbf{x}_i\mathbf{x}_i^H$ is the sample covariance matrix (SCM), and $\beta \in [0, 1]$ is a regularization parameter. It was shown in [5], [10] that setting a fixed value for β can be satisfactory for IPL (typically, we will use $\beta = .8$ in our experiments), but adaptive methods could also be found in [15]. The remainder of this section will now detail possible choices for the matrix distance d.

B. Existing IPL methods and corresponding distances

1) MLE-type IPL: Seminal IPL algorithms were driven by the assumption that the sample set follows a circularly symmetric complex Gaussian distribution, i.e. $\mathbf{x} = \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma}$ as in (3). The approximate MLE of $\mathbf{w}$ (assuming known $\mathbf{\Psi}$) then corresponds to the choice $\mathbf{\Sigma} = \mathbf{S}$ and the use of the Kullback-Leibler divergence

$$d^2_{\mathrm{KL}}(\mathbf{A}, \mathbf{B}) = \mathrm{tr}(\mathbf{B}^{-1}\mathbf{A}) + \log|\mathbf{B}\mathbf{A}^{-1}| - p. \tag{10}$$

After some manipulations, this MLE reduces to a COFI-PL problem (8), with

$$f^{\mathrm{KL}}_{\hat{\mathbf{\Sigma}}}(\mathbf{w}) = \mathbf{w}^{\mathsf{H}}(\hat{\mathbf{\Psi}}^{-1} \circ \hat{\mathbf{\Sigma}})\mathbf{w}. \tag{11}$$

Many works then motivated the use of other plug-in estimators within this formulation (cf. overview in [4]).

2) Least-squares IPL: The least-squares fitting estimator corresponds to the Euclidean distance

$$d^2_{\mathrm{E}}(\mathbf{A}, \mathbf{B}) = ||\mathbf{A} - \mathbf{B}||^2_{\mathrm{Fro}}. \tag{12}$$

With this distance COFI-PL then reduces to (8), with

$$f^{\mathrm{LS}}_{\hat{\mathbf{\Sigma}}}(\mathbf{w}) = -2\mathbf{w}^{\mathsf{H}}(\hat{\mathbf{\Psi}} \circ \hat{\mathbf{\Sigma}})\mathbf{w}, \tag{13}$$

which was shown to be of great practical interest in IPL [10]–[12], notably because this distance does not require any inversion of the covariance matrix plug-in.

C. Proposed new distances for IPL

Complex-valued covariance matrices belong to the space of Hermitian positive definite matrices $\mathcal{H}_p^{++}$. Endowing this space with a metric yields a Riemannian geometry for covariance matrices. An introduction to these concepts can be found in [13]. In this scope, many metrics can be envisioned, leading to various matrix distances, whose practical interests generally depend on the context and application. This paper will explore the use of these Riemannian distances in the COFI-PL formulation (6). We will focus on three prominent distances arising from various information geometry viewpoints [16].

1) Affine invariant (AI) distance: This distance corresponds notably to the Fisher-Rao information geometry of the Gaussian model [17]. The AI distance is defined as

$$d^2_{\mathrm{AI}}(\mathbf{A}, \mathbf{B}) = ||\log(\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2})||^2_{\mathrm{F}}. \tag{14}$$

It is characterized by the property

$$d^2_{\mathrm{AI}}(\mathbf{A}, \mathbf{B}) = d^2_{\mathrm{AI}}(\mathbf{X}\mathbf{A}\mathbf{X}^H, \mathbf{X}\mathbf{B}\mathbf{X}^H) \tag{15}$$

for any invertible matrix $\mathbf{X}$, i.e., this distance is invariant when both matrices are subject to the same affine transformation.

2) Log-Euclidean (LE) distance: This distance corresponds to an approximation of the affine invariant geometry, that yields some practical benefits in certain applications [18]. The LE distance is defined as

$$d^2_{\mathrm{LE}}(\mathbf{A}, \mathbf{B}) = ||\log(\mathbf{A}) - \log(\mathbf{B})||^2_{\mathrm{F}}. \tag{16}$$

3) Bures-Wasserstein (BW) distance: This distance is related to optimal transport of the Gaussian distribution, its corresponding geometry is studied in [19]. The BW distance is defined as

$$d^2_{\mathrm{BW}}(\mathbf{A},\mathbf{B}) = \mathrm{Tr}\{\mathbf{A}\} + \mathrm{Tr}\{\mathbf{B}\} - 2\mathrm{Tr}\{(\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2})^{1/2}\}. \tag{17}$$

IV. Riemannian optimization for COFI-PL

Given the distances from Section III-C, we now address the computation of the solution of (6). To do so, we resort to Riemannian optimization [20] on the torus $\mathbb{T}_p$ defined in (4).

A. Riemannian gradient descent

For the sake of conciseness, we will simply rely on the Riemannian gradient descent algorithm on $\mathbb{T}_p$ endowed with the Euclidean metric. More details concerning this choice can be found in [10]. In short, the iterates are produced as follows:

$$\mathbf{w}_{t+1} = \phi_{\mathbb{T}}(\mathbf{w}_t - \alpha_t \ \mathrm{grad} f(\mathbf{w}_t)), \tag{18}$$

In which

- $\phi_{\mathbb{T}}$ is the retraction operator that was defined after (3)
- The structure of the problem leads to the following definition of the Riemannian gradient:

$$\mathrm{grad} f(\mathbf{w}_t) = \nabla f(\mathbf{w}_t) - \mathfrak{Re}\{\nabla f(\mathbf{w}_t)^* \circ \mathbf{w}_t\} \circ \mathbf{w}_t, \tag{19}$$

  where $\nabla f(\mathbf{w}_t)$ is the Euclidean gradient of the objective function at $\mathbf{w}_t$.
- α_t is a step-size, that will be adaptively chosen at each step t according to the Armijo backtracking line search [21].

Given this algorithm, it remains to compute the Euclidean gradients $\nabla f(\mathbf{w}_t)$ for the objective in (7) constructed with the three distances presented in section III-C.

B. Euclidean gradients of $f^d_{\hat{\mathbf{\Sigma}}}$

The Euclidean gradient of a function $f : \mathbb{T}_p \to \mathbb{R}$ at point $\mathbf{w}_t$ is the unique vector defined as

$$\langle \nabla f(\mathbf{w}_t), \boldsymbol{\xi} \rangle = \mathrm{D}f(\mathbf{w}_t)[\boldsymbol{\xi}], \tag{20}$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean inner product, and where $\mathrm{D}f(\mathbf{w}_t)[\boldsymbol{\xi}]$ is the directional derivative of f with respect to $\mathbf{w}_t$ in direction $\boldsymbol{\xi}$. The objective function $f^d_{\hat{\mathbf{\Sigma}}}$ in (7) can be decomposed as

$$f^d_{\hat{\mathbf{\Sigma}}}(\mathbf{w}) = g(h(\mathbf{w})) \tag{21}$$

with

$$h : \mathbf{w} \in \mathbb{T}_p \mapsto \hat{\mathbf{\Psi}} \circ \mathbf{w}\mathbf{w}^{\mathsf{H}} \in \mathcal{H}^{++}_p \tag{22}$$

and

$$g : \mathbf{\Sigma} \in \mathcal{H}^{++}_p \mapsto d^2(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) \in \mathbb{R} \tag{23}$$

where d stands for any distance in (14) (16), or (17): we will use the notation g_{AI}, g_{LE}, and g_{BW} in accordance. The directional derivative f with respect to $\mathbf{w}$ computed along the direction $\xi_{\mathbf{w}} \in T_{\mathbf{w}}\mathbb{T}_p$ can be expressed through the following chain rule:

$$\mathrm{D}g(h(\mathbf{w}))[\xi_{\mathbf{w}}] = \mathrm{D}g(h(\mathbf{w}))[\mathrm{D}h(\mathbf{w})[\xi_{\mathbf{w}}]] \tag{24}$$

Algorithm 1 Riemannian gradient descent for (6)

In: $\hat{\mathbf{\Sigma}} \in \mathcal{H}^{++}_p$, $d \in \{d_{\mathrm{AI}},\ d_{\mathrm{LE}},\ d_{\mathrm{BW}}\}$, $\mathbf{w}_0 \in \mathbb{T}_p$, $\alpha_0 \in \mathbb{R}^*_+$
Set $t = 1$
repeat
  Compute Euclidean gradient with Proposition 1-2-3
  Compute Riemannian gradient with (19)
  Update α_t with Armijo backtracking [21]
  Produce iterate $\mathbf{w}_t$ with (18)
  $t = t + 1$
until convergence
Out: $\hat{\mathbf{w}} = \mathbf{w}_{\mathrm{end}} \in \mathbb{T}_p$

We thus need to compute two quantities:

- $\mathrm{D}h(\mathbf{w})[\xi_{\mathbf{w}}]$ for $\xi_{\mathbf{w}} \in T_{\mathbf{w}}\mathbb{T}_p$ is directly obtained from the expression of h in (22) as

$$\mathrm{D}h(\mathbf{w})[\xi_{\mathbf{w}}] = \hat{\mathbf{\Psi}} \circ \xi_{\mathbf{w}}\mathbf{w}^{\mathsf{H}} + \hat{\mathbf{\Psi}} \circ \mathbf{w}\xi^{\mathsf{H}}_{\mathbf{w}}$$

- $\mathrm{D}g(\mathbf{\Sigma})[\xi_{\mathbf{\Sigma}}]$ for $\xi_{\mathbf{\Sigma}} \in T_{\mathbf{\Sigma}}\mathcal{H}^{++}_p$ depends on the chosen distance, and the following results can be obtained from [22] and [14]:

$$\mathrm{D}g_{\mathrm{AI}}(\mathbf{\Sigma})[\xi_{\mathbf{\Sigma}}] = 2\mathrm{Tr}\{\hat{\mathbf{\Sigma}}^{1/2}\boldsymbol{\xi}_{\mathbf{\Sigma}}\hat{\mathbf{\Sigma}}^{1/2}\log(\hat{\mathbf{\Sigma}}^{-1/2}\mathbf{\Sigma}\hat{\mathbf{\Sigma}}^{-1/2})\}$$

$$\mathrm{D}g_{\mathrm{LE}}(\mathbf{\Sigma})[\xi_{\mathbf{\Sigma}}] = 2\mathrm{Tr}\{[\mathbf{\Sigma}^{-1}\log(\mathbf{\Sigma}) - d\log(\mathbf{\Sigma})[\log(\hat{\mathbf{\Sigma}})]]\boldsymbol{\xi}_{\mathbf{\Sigma}}\}$$

$$\mathrm{D}g_{\mathrm{BW}}(\mathbf{\Sigma})[\xi_{\mathbf{\Sigma}}] = \mathrm{Tr}\{[\mathbf{I} - \hat{\mathbf{\Sigma}}^{1/2}(\hat{\mathbf{\Sigma}}^{1/2}\mathbf{\Sigma}\hat{\mathbf{\Sigma}}^{1/2})^{-1/2}\hat{\mathbf{\Sigma}}^{1/2}]\boldsymbol{\xi}_{\mathbf{\Sigma}}\}$$

Combining the above directional derivatives, and by identification with (20), we obtain the following gradients for each objective function.

***Proposition** 1:* The Euclidean gradient of $f^{\mathrm{AI}}_{\hat{\mathbf{\Sigma}}}$ at point $\mathbf{w}_t$ is defined as

$$\nabla f^{\mathrm{AI}}_{\hat{\mathbf{\Sigma}}}(\mathbf{w}_t) = 4 \cdot [\hat{\mathbf{\Sigma}}^{1/2}\log(\hat{\mathbf{\Sigma}}^{-1/2}h(\mathbf{w}_t)\hat{\mathbf{\Sigma}}^{-1/2})\hat{\mathbf{\Sigma}}^{1/2}]\hat{\mathbf{\Psi}}\mathbf{w}_t \tag{25}$$

***Proposition** 2:* The Euclidean gradient of $f^{\mathrm{LE}}_{\hat{\mathbf{\Sigma}}}$ at point $\mathbf{w}_t$ is defined as

$$\nabla f^{\mathrm{LE}}_{\hat{\mathbf{\Sigma}}}(\mathbf{w}_t) = 4 \cdot \mathrm{D}\log(h(\mathbf{w}_t))[\log(h(\mathbf{w}_t)) - \log(\hat{\mathbf{\Sigma}})]\hat{\mathbf{\Psi}}\mathbf{w}_t \tag{26}$$

and the quantity $\mathrm{D}\log(\mathbf{\Sigma})[\xi_{\mathbf{\Sigma}}]$ can be evaluated numerically thanks to the relation [23]

$$\log\left(\begin{bmatrix}\mathbf{\Sigma} & \boldsymbol{\xi}_{\mathbf{\Sigma}} \\ \mathbf{0} & \mathbf{\Sigma}\end{bmatrix}\right) = \begin{bmatrix}\log(\mathbf{\Sigma}) & \mathrm{D}\log(\mathbf{\Sigma})[\xi_{\mathbf{\Sigma}}] \\ \mathbf{0} & \log(\mathbf{\Sigma})\end{bmatrix} \tag{27}$$

***Proposition** 3:* The Euclidean gradient of $f^{\mathrm{BW}}_{\hat{\mathbf{\Sigma}}}$ at point $\mathbf{w}_t$ is defined as

$$\nabla f^{\mathrm{BW}}_{\hat{\mathbf{\Sigma}}}(\mathbf{w}) = 2 \cdot [\mathbf{I} - \hat{\mathbf{\Sigma}}^{1/2}(\hat{\mathbf{\Sigma}}^{1/2}h(\mathbf{w}_t)\hat{\mathbf{\Sigma}}^{1/2})^{-1/2}\hat{\mathbf{\Sigma}}^{1/2}]\hat{\mathbf{\Psi}}\mathbf{w}_t \tag{28}$$

The propositions just presented allow to compute the Riemannian gradient for all the proposed COFI-PL formulations using (19) and thus, to perform the Riemannian gradient descent using the iteration (18). The resulting algorithm is summarized in the box Algorithm 1.

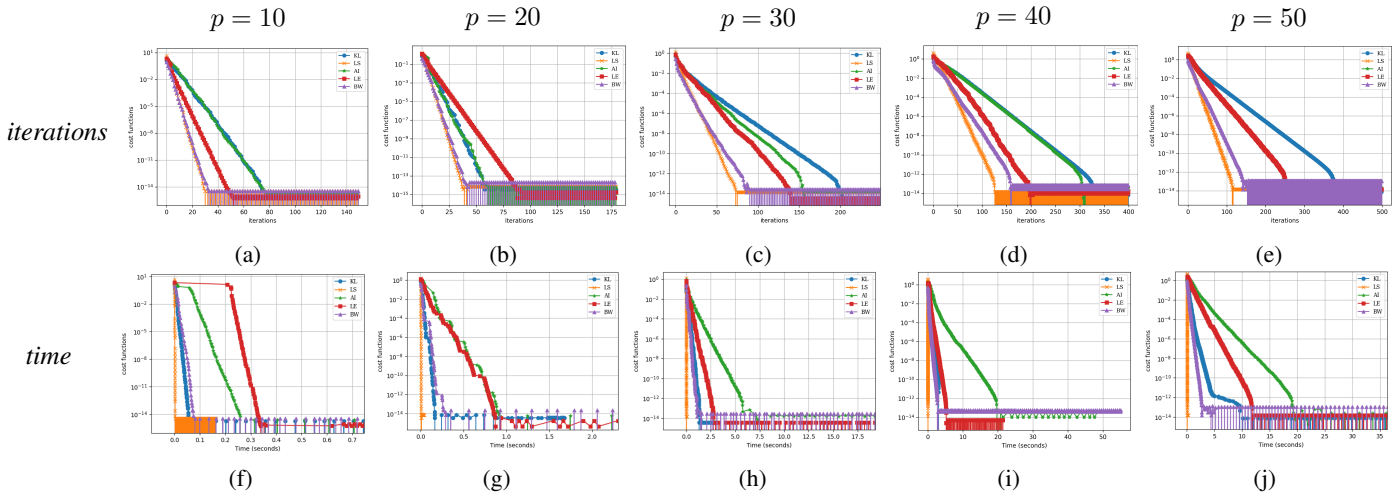

Fig. 1: Cost function optimizations (log-scale) for varying sample sizes ($p = 10,\ 20,\ 30,\ 40,\ 50$) for KL, Euclidean (LS), AI, LE and BW distances using the Armijo-style Steepest Riemannian Gradient Descent optimisation w.r.t. iterations (a-e) and runtime (f-j). The parameters were set to $n = 2p,\ \rho = 0.7$.

V. Numerical Results

A. Methods

In this section, we compare the merits of the different distances that can be used in (6). The algorithm developed in this paper allows us to use the Affine-Invariant (AI) distance in (16), the Log-Euclidean (LE) distance in (14), or the Bures-Wasserstein (BW) distance in (17). We also include in this study the Kullback-Leibler (KL) divergence in (11) and the Euclidean distance (LS) in (13), that were studied in [10].

B. Setup parameters

The comparisons are conducted on simulated data with varying underlying parameters. For a chosen dimension p the covariance matrix $\boldsymbol{\Sigma}$ is constructed as in (3), where $\boldsymbol{\Psi}$ is set as a Toeplitz matrix of correlation coefficient ρ, i.e.,

$$[\boldsymbol{\Psi}]_{q\ell} = \rho^{|q-\ell|}. \tag{29}$$

Phase differences vary linearly between 0 and 2 radians, i.e. $\Delta_{i+1,i} = \theta_{i+1} - \theta_i = 2/p$ rad. The samples are simulated through a Gaussian distribution as $\boldsymbol{x}_i \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$. Then the RSCM as defined in (9), is used as plug-in estimator in the COFI-PL formulation (6).

C. Comparison of distances in terms of computational load

Because we rely on the same generic algorithm (Riemannian Gradient Descent with Armijo backtracking rule [21]), we can fairly compare the complexity by investigating the number of required iterations, the computation time, as well as the computational complexity associated with each distance. For varying dimensions, the value of the cost function is observed throughout the optimisation process w.r.t. its iterations (Fig. 1a-1e) and its runtime (Fig. 1f-1j). The LS distance was always observed to lead to the fastest convergence. For smaller sample sizes, the LE distance proved to be the most time consuming, taking almost twice as much time to minimize compared to the second slowest algorithm. However, with increasing sample sizes the AI distance starts having the largest runtime. For a more complete understanding of this phenomenon, the computational complexity of each distance must be discussed. Based on the expressions of the gradients previously presented, the Euclidean gradient has the lowest computational complexity, namely $\boldsymbol{O}(n^2)$, compared to the gradients of the other distances which present a computational complexity of $\boldsymbol{O}(n^3)$. A particular note must be made for the LE and AI distances. While in the $\boldsymbol{O}(n^3)$ order of complexity, both distances require multiple operations with this same complexity order to be computed at each iteration, namely, for the LE distance, the matrix logarithm, its derivative, and the inverse of the matrix, while for the AI distance, an Eigenvalue Decomposition is carried out in order to evaluate the inverse square root of a matrix and its logarithm. For both cases, this leads to a high complexity overhead, with a higher overhead for the LE which is visible for smaller sample dimensions. However, AI involves more complex matrix operations that start to affect runtime with increasing matrix dimensions.

D. Comparison of distances in terms of estimation accuracy

Methods are compared by computing the Mean Squared Error (MSE) on the phase of the last date (i.e. $[w]_p$), that is evaluated using 1000 Monte-Carlo trials. For each distance, the optimisation process was conducted for 3000 iterations or until the tolerance fell below the $\mathcal{O}(10^{-4})$ order of magnitude. In this scenario, we set $p = 10$ with varying n or varying ρ:

1) $n \in [p, 4p]$, for $\rho^p = 0.7$ in Fig. 2.
2) $\rho^p \in [0.5, 0.9]$, for $n = p + 2$ in Fig. 3.

We also notice that no particular trend differences were observed when changing the dimension p. The results show that the MSE decreases for algorithms with either increasing n (Fig. 2), or increasing temporal coherence (Fig. 3), which was to be expected. We also notice a similar trend for all

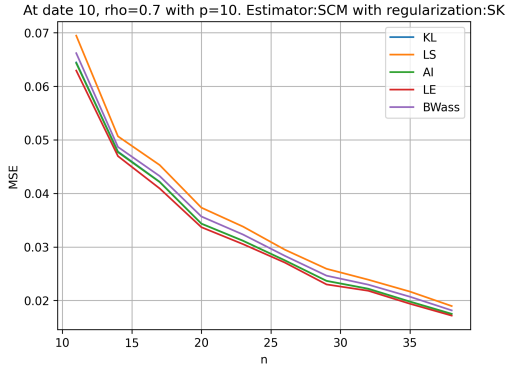

Fig. 2: MSE of the last date for KL, Euclidean (LS), AI, LE and BW distances w.r.t. $n \in [p, 4p]$ for $\rho = 0.7$ and $p = 10$

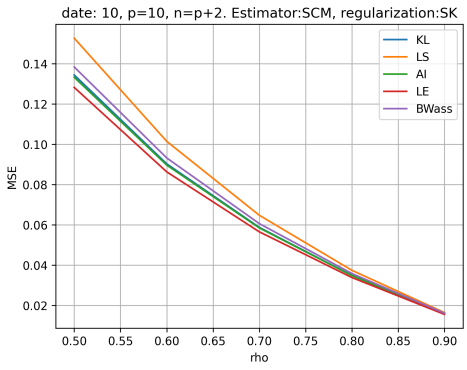

Fig. 3: MSE of the last date for KL, Euclidean (LS), AI, LE and BW distances w.r.t. $\rho \in [0.5, 0.9]$ for $p = 10$ and $n = p+2$

methods, with no crossing in terms of performance. Overall, LE provides the best accuracy for all sample dimensions and all coherence values. However, since the performances tend towards similar values at high coherence, the computational costs and time constraints could be taken into consideration when choosing which distance to use.

VI. Conclusions

In multitemporal interferometric SAR, this paper investigated the use of Riemannian distances (affine invariant, log-Euclidean, and Bures-Wasserstein) on covariance matrices in order to estimate phase differences between the images. The resolution of the corresponding covariance fitting problems was addressed using Riemannian optimizations. Simulations providing comparisons between the different choices of distances evidenced a trade-off between computational complexity and estimation accuracy. An adaptive selection of the matrix distance with respect to the estimated coherence of the pixel patch could thus bring the best of both worlds.

References

[1] Dinh Ho Tong Minh and Stefano Tebaldini, "Interferometric phase linking: Algorithm, application, and perspective," *IEEE Geoscience and Remote Sensing Magazine*, vol. 11, pp. 46–62, 09 2023.

[2] Andrea Monti Guarnieri 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] Simon Zwieback, "Cheap, valid regularizers for improved interferometric phase linking," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–4, 2022.

[6] Michael Schmitt, Johannes L Schönberger, and Uwe Stilla, "Adaptive covariance matrix estimation for multi-baseline InSAR data stacks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 6807–6817, 2014.

[7] Yuanyuan Wang and Xiao Xiang Zhu, "Robust estimators for multipass sar interferometry," *IEEE Transactions On Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 6807–6817, 2016.

[8] Phan Viet Hoa Vu, Arnaud Breloy, Frédéric Brigui, Yajing Yan, and Guillaume Ginolhac, "Robust Phase Linking in InSAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–11, 2023.

[9] 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.

[10] Phan Viet Hoa Vu, Arnaud Breloy, Frédéric Brigui, Yajing Yan, and Guillaume Ginolhac, "Covariance fitting interferometric phase linking: Modular framework and optimization algorithms," *arXiv preprint arXiv:2403.08646*, 2024.

[11] Phan Viet Hoa Vu, Arnaud Breloy, Frédéric Brigui, Yajing Yan, and Guillaume Ginolhac, "Covariance fitting based InSAR Phase Linking," in *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2023, pp. 8234–8237.

[12] Yusong Bai, Jian Kang, Xiang Ding, Anping Zhang, Zhe Zhang, and Naoto Yokoya, "LaMIE: Large-Dimensional Multipass InSAR Phase Estimation for Distributed Scatterers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.

[13] Florent Bouchard, Arnaud Breloy, Antoine Collas, Alexandre Renaux, and Guillaume Ginolhac, "The fisher–rao geometry of ces distributions," in *Elliptically Symmetric Distributions in Signal Processing and Machine Learning*, pp. 37–77. Springer, 2024.

[14] Yann Thanwerdas and Xavier Pennec, "The geometry of mixed-Euclidean metrics on symmetric positive definite matrices," *Differential Geometry and its Applications*, vol. 81, pp. 101867, 2022.

[15] Esa Ollila and Arnaud Breloy, "Regularized tapered sample covariance matrix," *IEEE Transactions on Signal Processing*, vol. 70, pp. 2306–2320, 2022.

[16] Yann Thanwerdas and Xavier Pennec, "O(n)-invariant riemannian metrics on spd matrices," *Linear Algebra and its Applications*, vol. 661, pp. 163–201, 2023.

[17] Arnaud Breloy, Guillaume Ginolhac, Alexandre Renaux, and Florent Bouchard, "Intrinsic Cramér–Rao bounds for scatter and shape matrices estimation in CES distributions," *IEEE Signal Processing Letters*, vol. 26, no. 2, pp. 262–266, 2018.

[18] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache, "Log-Euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 411–421, 2006.

[19] Rajendra Bhatia, Tanvi Jain, and Yongdo Lim, "On the bures–wasserstein distance between positive definite matrices," *Expositiones Mathematicae*, vol. 37, no. 2, pp. 165–191, 2019.

[20] Nicolas Boumal, *An introduction to optimization on smooth manifolds*, Cambridge University Press, 2023.

[21] Larry Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of mathematics*, vol. 16, no. 1, pp. 1–3, 1966.

[22] Maher Moakher, "A differential geometric approach to the geometric mean of symmetric positive-definite matrices," *SIAM journal on matrix analysis and applications*, vol. 26, no. 3, pp. 735–747, 2005.

[23] Nicholas J Higham, *Functions of matrices: theory and computation*, SIAM, 2008.