SEQUENTIAL ESTIMATOR FOR DISTRIBUTED SCATTERER INTERFEROMETRY

Homa Ansari [(1)], Francesco De Zan [(1)], Nico Adam [(1)], Kanika Goel [(1)], Richard Bamler [(1,2)]

(1) Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling, Germany
(2) Chair of Remote Sensing Technology, Technical University of Munich (TUM), Munich, Germany
Email: homa.ansari@dlr.de

ABSTRACT

The launch of the wide-swath SAR missions with short repeat-pass cycles, such as Sentinel-1, will soon provide an unprecedented large InSAR data archive. Time-series analysis on the rapidly growing data will thus become computationally demanding for a systematic monitoring of earth surface deformation. As the state-of-the-art approach in differential InSAR time-series analysis, the distributed scatterer interferometric (DSI) techniques shall adapt agile processing schemes to deal with the emerging big data; an aspect to which limited attention has been dedicated. In this contribution, a sequential DSI scheme is proposed to address this demand. Based on SAR data reduction, the scheme allows for batch processing of the large data stacks while preserving the performance close to the Cramér-Rao Lower Bound. The performance of the proposed sequential estimator is compared to the current DSI algorithms under two contradicting coherence scenarios. The application of the proposed sequential estimator to stacks of Sentinel-1 data is ongoing.

Index Terms— Big InSAR data, Distributed Scatterer Interferometry, Dimensionality Reduction, Low-Rank Approximation, Performance Analysis, Robust Estimation

1. INTRODUCTION

Distributed Scatterer Interferometry (DSI) is a framework for retrieving geophysical signal from time-series of SAR data. In this framework a number of algorithms have been proposed for estimation of the deformation phase corresponding to the geophysical signal of interest. Examples are: Small Baseline Subset Approach (SBAS) [1], SqueeSAR [2] and CAESAR [3]. Although proved to be powerful and applicable in time-series analysis, the proposed algorithms are computationally expensive and do not allow near-real-time processing of the data stacks.

The advent of wide swath SAR missions with short repeat-pass time and global coverage, such as Sentinel-1, will introduce large SAR data archives. This Big SAR data calls for more agile stacking techniques with *sequential* processing capabilities. In the context of sequential processing, the Kalman filter approach is a trivial solution.

It is however based on an explicit definition of a dynamic system (in here, a geophysical deformation model) and prone to misspecification of such a system. In [4] the authors proposed an alternative scheme tailored to a long-term coherence scenario. The algorithm is here referred to as *virtual estimator*. Inspired by the idea of the virtual estimator, the aim here is on establishing a generic precise sequential DSI scheme; with performance close to the theoretical Cramér-Rao Lower Bound (CRLB) [5]. The performance of the proposed sequential estimator is compared to the available DSI algorithms through contradicting simulation cases. Experiments with Sentinel-1 and large stack of TerraSAR-X data has been performed and is ongoing.

2. DISTRIBUTED SCATTERER INTERFEROMETRY

Among the available DS estimators, the SqueeSAR approach provides the Maximum Likelihood Estimation (MLE) of the deformation phase. Under the assumption of complex circular Gaussian statistics of the SAR measurements, the MLE is asymptotically the closest estimator to the CRLB in the estimation of the deformation phase. However, as it will be shown, the performance of MLE is affected by the assumed stochastic model; i.e. the estimated coherence. If coherence estimation is biased, the MLE cannot approach its asymptotic performance. This will often be the case due to the limited size of the ensemble used for coherence estimation and/or presence of outliers in the ensemble [6].

2.1. The impact of decorrelation model on deformation phase retrieval

In order to show the impact of the stochastic model on the different estimators, two contradicting coherence models have been considered. The first manifests a pure exponential decorrelation between the interferometric pairs, i.e.:

$$\Gamma_{i,k} = \gamma_0 \exp\left(\frac{-t_{i,k}}{\tau_0}\right), \tag{1}$$

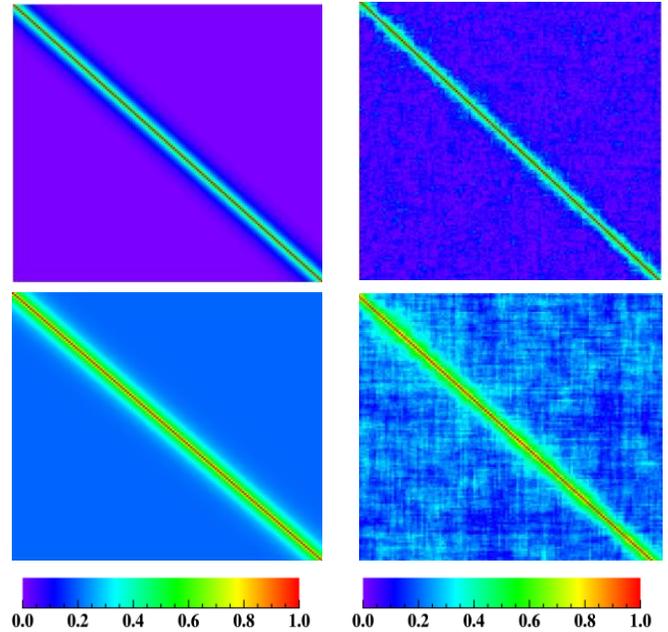while the second reveals a residual coherence even for large temporal baselines, i.e. [7]:

$$\Gamma_{i,k} = (\gamma_0 - \gamma_\infty) \exp\left(\frac{-t_{i,k}}{\tau_0}\right) + \gamma_\infty. \tag{2}$$
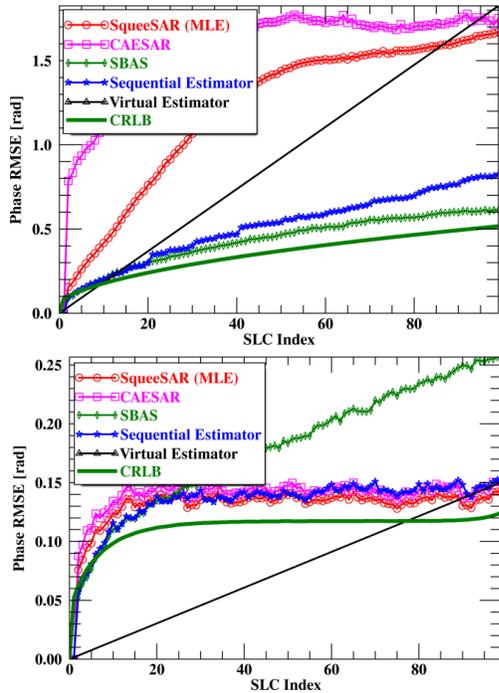
Figure 1: Performance of DS phase estimation schemes relevant to the different coherence scenarios: top: exponential decorrelation, bottom: long-term coherence. Here, CAESAR exploits the largest principle component; the virtual estimator is only used for the phase retrieval of the last SLC; and the SBAS-like algorithm is bounded to short temporal baselines of less than 60 days (equivalent to lag 10 interferograms). Note that the performance of MLE is degraded with the biased estimation of the coherence (see Fig. 2), while the sequential estimator maintains a balanced performance in both cases.

Figure 2: Coherence model as the second order stochastic of the simulated data stack; top row: exponential decorrelation, bottom row: long-term coherence, left column: simulated coherence matrix, right column: estimated coherence matrix for an ensemble of 200 statistically homogeneous pixels. Note the severe estimation bias for coherence close to zero.

In these formulations, $\Gamma_{i,k}$ is the coherence between the ith and k^{th} SLC of the stack, γ_0 and γ_∞ respectively indicate the initial and residual coherence, $t_{i,k}$ stands for the temporal baseline and τ_0 is the time constant of the decorrelation process.

Assuming a circular complex Gaussian statistics , a stack of 100 SLCs each containing 200 statistically homogenous pixels (looks) is simulated, using the aforementioned decorrelation models: the temporal sampling interval is set to 6 days (as will be the case for the Sentinel-1) and the deformation phase is set to zero, the decorrelation model parameters are provided in Table 1. The objective is to retrieve the phase history for each simulated SLC with the available DS estimators. In phase retrieval, CAESAR exploits the largest principle component of the coherence matrix; the virtual estimator is only used for phase estimation of the last SLC in the stack; and the SBAS-like algorithm is bounded to interferometric pairs with short temporal baselines of less than 60 days (for different simulation scenarios the corresponding coherence threshold can be calculated using Eq. 1 and 2). To assess the performance of different schemes, the experiment is repeated 500 times and the Root Mean Square Error (RMSE) of the estimated compared to the simulated phase is reported. Table 1 summarizes the simulation cases as well as

the RMSE of phase estimation for the last SLC in the simulated stack.

Fig. 1 depicts the performance of different estimators compared to the theoretical CRLB for the two considered coherence scenarios. Note that the CRLB is calculated with the theoretical coherence given by Eq. 1 and 2 while the different estimators exploit an estimation of the coherence. The coherence is estimated using an ensemble of 200 statistically homogenous pixels. The simulated and estimated coherence matrices are provided in Fig 2. As evident from Fig 1, the performance of estimators strongly depends on the coherence model. In case of exponential decorrelation, although MLE is expected to be the closest to CRLB, the SBAS-like approach outperforms this estimator. The reason lies in the biased estimation of coherence (see Fig. 2). The coherence is overestimated for coherence values close to zero, thus leading to biased results of MLE compared to SBAS. The latter approach simply discards the interferometric pairs with long temporal baselines and safeguards against the inclusion of interferometric pairs with overestimated coherence (which in reality are noisy interferometric phases baring no coherent signal). This observation highlights the importance of robust estimation schemes, such as the M-estimator proposed in [6], not only for safeguarding against the outliers, as discussed in [6], but also for modification of the stochastic model relative to estimation residuals.

Table 1: Performance of the different phase triangulation schemes compared to the CRLB; reported is the RMSE of the estimated and the simulated deformation phase for the last SLC of the stack. The performance of the proposed sequential estimator is indicated with block letters.

	Model	Exponential Decay	Long-Term Coherence
Coherence Scenario	Equation	1	2
	Simulation Parameters	$\gamma_0 = 0.6$ $\tau_0 = 27$ days	$\gamma_0 = 0.8, \gamma_\infty = 0.2$ $\tau_0 = 27$ days
RMSE phase estimation [rad]	SqueeSAR	1.67	0.14
	CAESAR	1.71	0.15
	SBAS	0.62	0.26
	Virtual Est.	1.82	0.16
	Sequential Est.	**0.83**	**0.15**
	CRLB	0.52	0.12

Examining the long-term coherence scenario, the performance of SBAS severely degrades compared to MLE. This performance loss is due to discarding the low-coherence interferograms ($\gamma_\infty = 0.2$). Such an observation reveals the importance of inclusion of even low-coherent (but non-zero) interferometric pairs in deformation phase estimation.
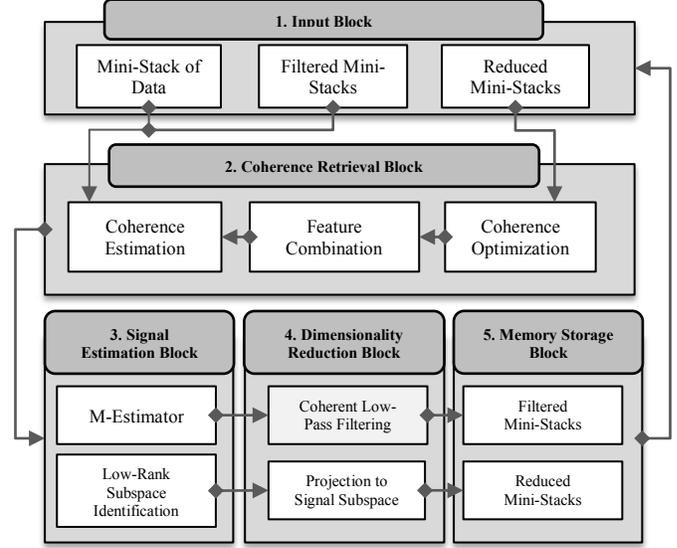
3. THE PROPOSED SEQUENTIAL DS ESTIMATOR

In the development of a sequential estimator two contradicting goals shall be preserved. On one hand, the aim is on batch processing of data stream without acquiring and exploiting the entire stack, which in itself reduces the computational burden. On the other hand, the batch processing shall not cause performance loss due to neglecting (even small) coherence among the archived and streamed data. From the previous section, it is evident that even neglecting low level coherence among the data leads to dramatic performance loss.

Inspired by the virtual estimator [4], here a sequential scheme is established to retrieve the correlation between the streamed and the processed data. The method is based on acquisition of a limited number of SLCs, hereafter referred to as *mini-stacks*, batch processing of the streamed mini-stack and passing its information content to the future processing chain via projection of the data to its low-rank signal subspace. The stored extracted information is used to retrieve its coherence with the future data stream, so that the data history is not neglected and is to some extent retrieved for phase estimation. The flowchart of the established method is provided in Fig. 3. A selection of the processing blocks is explained in the following:

3.1. Dimensionality reduction of SAR data stacks
In order to make the information content of each batch available to the future data stream, the streamed mini-stack is reduced to its lower-dimension signal subspace. The compressed data is stored in the memory and utilized for coherence retrieval. The data compression is sought in two

Figure 3: Algorithmic flow of the proposed DS sequential estimator

steps:

3.1.1. *Temporal Coherent Low Pass Filtering (CLPF):*
The hidden geophysical signal in the data stream can be interpreted as a low-pass component and retrieved by phase triangulation schemes. In our proposed approach this signal component is estimated using the robust M-Estimator suggested in [6]. The M-estimator not only allows for suppression of the outliers but also safeguards against the inconsistent interferometric pairs by down-weighing their contribution in the triangulation.

As the first level of data reduction, the mini-stack of streamed SLCs is coherently filtered by the estimated phase [4]:

$$Z_{CLPF}(l) = \frac{1}{S} \sum_{k=0}^{S} Z_k(l) \exp(-j \, \hat{\phi}_k), \qquad (3)$$

here S is the size of the mini-stack, l stands for the lth look of the homogenous pixels at each SLC, k refers to the SLC index in the data batch, Z_k and $\hat{\phi}_k$ indicate the complex valued pixel and the estimated phase of the corresponding SLC.

3.1.2. *Projection to Signal Subspace*
To assure maximum information extraction from the mini-stacks, we extend the data reduction beyond the low-pass filtering. The extension is carried out by identification of the low-dimensional subspace which spans the underlying meaningful signal of the mini-stack and separates it from the noise subspace.

In its simplest form, signal subspace identification can be carried out by Principle Component Analysis (PCA). Similar to CAESAR, the PCA can decompose the coherence matrix into the scattering features [3]:

$$\gamma_S = \sum_{i=0}^{S} \lambda_i \, \boldsymbol{u}_i \boldsymbol{u}_i^H, \qquad (4)$$

where γ_S is the complex coherence matrix of the batch and complex vectors $\boldsymbol{u}_i$ are its decomposed features. Dependent on the significance of the features to the data content, the most significant principle components may be chosen for data reduction. The number of representative components can be related to the entropy of the data. The mini-stack is then projected to the signal subspace by:

$$Z_{\text{RED,i}}(l) = \sum_{k=0}^{s} Z_k(l) u_{k,i}. \tag{5}$$

The subscripts i and k index the principle component and the SLC number respectively.

Although utilized for the first demonstration of the sequential algorithm, PCA is not the optimum scheme for subspace identification. The PCA seeks the optimum subspace representation in the L2 norm sense and is thus highly sensitive to outliers. This limitation can be tackled by the use of robust techniques such as robust PCA [9] that provide a robust estimation of the optimum low-rank signal subspace.

3.2. Coherence retrieval

The stored compressed data is exploited to retrieve the coherence between the streamed mini-stack and the unavailable data archive. The stored Low-pass filtered components are directly integrated to the sequential processing. The integration of the higher order information content, captured by the projection of the data to the estimated signal subspace, shall however be handled with more care as their direct inclusion may introduce noise in the sequential process. The relevance of the higher order information to the streamed data is assured in the *coherence optimization* step. In this step coherence of the reduced-mini-stacks with respect to the streamed data is maximized. This step follows the formulation of eigenvalue problem proposed for polarimetric interferometry [8]:

$$\Gamma_{22}^{-1} \Gamma_{12}^{T} \Gamma_{11}^{-1} \Gamma_{12} \ \boldsymbol{w} = \lambda \boldsymbol{w}, \tag{6}$$

here, Γ_{22} and Γ_{11} are the estimated coherence of the archived reduced data i.e. $Z_{\text{RED},i}$ and the streamed data i.e. Z, respectively, while Γ_{12} is the coherence between the two mentioned data sets. The coherence is maximized by the optimal weights in $\boldsymbol{w}$. The reduced data is therefore linearly-combined given these weights, i.e.:

$$Z_{\text{Combined}}(l) = \sum_{i=0}^{NC} w_i Z_{\text{RED,i}}(l), \tag{7}$$

with NC as the size of the reduced data. The streamed data is finally appended by the LP features as well as the linear-combination of the higher order features. Estimating the coherence with the appended features, the coherence between the streamed data and the unavailable data history is partially retrieved. This estimated coherence is the input to the signal estimation block.

3.3. Performance of Sequential DS Estimator

Performance of the proposed sequential DS estimator is provided in Fig. 1 and Table 1. As apparent form the comparisons, the proposed algorithm maintains performance in the two contradicting decorrelation scenarios as opposed to the non-sequential DS schemes. The RMSE is fairly close to the CRLB and is not degraded by the coherence scenario.

4. CONCLUDING REMARKS

A sequential estimation scheme has been proposed for exploitation of distributed scatterers in the multi-temporal InSAR data stacks. The scheme enables near real-time processing of InSAR time series, with performance close to CRLB. The performance of the estimator has been compared to the available DS techniques under two idealized contradicting coherence scenarios.

Future work focuses on: investigation of dimensionality reduction techniques for multi-temporal SAR data; improving the robustness of the algorithms involved to safeguard against the impact of outliers; sequential detection of the statistically homogenous regions; and performance assessment of the technique by application to large stacks of Sentinel-1 data.

5. REFERENCES

[1] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375–2383, Nov. 2002.

[2] A. Ferretti *et al.*, "A new algorithm for processing interferometric data stacks: SqueeSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3460–3470, Sep. 2011.

[3] G. Fornaro, S. Verde, D. Reale, and A. Pauciullo, "CAESAR: An approach based on covariance matrix decomposition to improve Multi-baseline multi-temporal interferometric SAR processing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2050–2065, Apr. 2015.

[4] F. De Zan, and P. López-Dekker. "SAR Image Stacking for the Exploitation of long-term coherent Targets." *Geoscience and Remote Sensing Letters, IEEE* 8.3 (2011): 502-506.

[5] A. Monti Guarnieri and S. Tebaldini, "Hybrid Cramér Rao bound for crustal displacement field estimators in SAR interferometry," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 1012–1015, Dec. 2007.

[6] Y. Wang, and X. X. Zhu. "Robust Estimators for Multipass SAR Interferometry." *IEEE Trans. Geosci. Remote Sens.*, 2015.

[7] A. Parizzi, X. Y. Cong, and M. Eineder, "First results from multi-frequency interferometry—A comparison of different decorrelation time constants at X, C and L-band," in *Proc. Fringe Workshop*, 2009.

[8] S. Cloude and K. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1551–1565, Sep. 1998.

[9] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization." *Advances in neural information processing systems,* pp. 2080-2088, 2009.