M-NL: Robust NL-Means Approach for PolSAR Images Denoising

Gordana Draškovic, *Student Member, IEEE*, Frédéric Pascal, *Senior Member, IEEE*, and Florence Tupin, *Senior Member, IEEE*

***Abstract*— This letter proposes a new method for polarimetric synthetic aperture radar (PolSAR) denoising. More precisely, it seeks to address a new statistical approach for weights computation in nonlocal (NL) approaches. The aim is to present a simple criterion using *M*-estimators and to detect similar pixels in an image. A binary hypothesis test is used to select similar pixels which will be used for covariance matrix estimation together with associated weights. The method is then compared with an advanced state-of-the-art PolSAR denoising method named NL-SAR. The filter performances are measured by a set of different indicators, including relative errors on incoherent target decomposition parameters, coherences, polarimetric signatures, and edge preservation on a set of simulated PolSAR images. Finally, results for RADARSAT-2 PolSAR data are presented.**

***Index Terms*— Detection, *M*-estimators, nonlocal (NL) means, polarimetric synthetic aperture radar (PolSAR), Wishart distribution.**

I. Introduction

In THE past decades, there has been a growing interest for polarimetric synthetic aperture radar (PolSAR) images and their use for terrain classification, target detection, and so on. Speckle in PolSAR data significantly degrades the image quality as well as the application performances. In PolSAR images, each pixel is given by a complex-valued matrix (or vector) formed out of backscattered signals in different combinations of the linear received and transmitted polarizations. Since the scatterers are distributed and due to the coherency of PolSAR systems, this matrix has a random nature and is referred to as speckle noise. Therefore, in order to determine the physical parameters of interest, a speckle reduction step is usually applied, aiming at reducing the parameter fluctuations. Different methods have been proposed to reduce speckle in PolSAR data [3]. Recent approaches of image processing and computer vision rely on patch-based comparison to select similar samples. The idea was born with the first works of Buades *et al.* [4], [5] and then extended and adapted to (Pol)SAR images by several authors, such as in [6] and [7]. In these methods, the covariance matrices (CMs) of scattering vectors (SVs) are estimated using carefully selected samples in order to reduce the speckle. The SVs are usually modeled by complex circular Gaussian vectors and completely characterized by their CM. The CM, unknown in practice, is then obtained by averaging, i.e., using the well-known sample CM (SCM), which, in the Gaussian context, is known to be the maximum likelihood (ML)-estimator (MLE) and Wishart-distributed.

Nevertheless, the Gaussian distribution is not always a good approximation for PolSAR data, since, for instance, this model is not able to describe textured scenarios. Therefore, it is necessary to increase the complexity of the underlying distribution model. This has been successfully done by introducing the so-called spherically invariant random vectors (SIRVs) [8]. The SIRV displays an observation as a product of a Gaussian random process with the square root of a nonnegative random scalar variable that contains the information about the texture variation. Consequently, the SIRVs encompass a wide class of well-known distributions, such as t-distribution, K-distribution, or inverse Gaussian texture distribution. In a non-Gaussian environment, the SCM may not provide accurate results for the CM estimate. For this reason, *M*-estimators have been introduced and broadly employed and analyzed in robust statistics and signal processing [9]–[11]. *M*-estimators are a wide class of scatter matrix (SM)[1] estimators robust to the data model. These estimators are given by fixed-point equations which makes the direct analysis of their statistical properties very difficult. Nevertheless, it has been recently shown that the behavior of *M*-estimators can be well approximated by Wishart distribution [12], [13] which is of great importance in various applications.

In this letter, thanks to the new properties of *M*-estimators, we introduce a new NL-means algorithm to estimate CM in PolSAR data. Then, we evaluate the performances of this method and compare it with the NL-SAR method. NL-SAR is a state-of-the-art method composed of several steps: preestimation, weights computation, bias reduction, and best estimate selection. In this letter, the focus is on the two first steps.

1) *Preestimation:* Instead of using a Gaussian kernel, we propose to use an *M*-estimator in order to preestimate matrices in each pixel and compare patches.
2) *Weight Computation:* In this step, instead of learning kernels to weights similarities, we first select similar pixels based on the statistics of a robust similarity test and, then, compute weights using only selected pixels.

Manuscript received September 2, 2018; revised October 16, 2018 and November 11, 2018; accepted November 29, 2018. Date of publication January 11, 2019; date of current version May 21, 2019. This work was supported by the Labex DigiCosme/Digiteo. *(Corresponding author: Frédéric Pascal.)*

G. Draškovic and F. Pascal are with L2S, CentraleSupélec, CNRS, Université Paris-Sud, F-91192 Gif-sur-Yvette, France (e-mail: gordana.draskovic@l2s.centralesupelec.fr; frederic.pascal@l2s.centralesupelec.fr).

F. Tupin is with LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France (e-mail: florence.tupin@telecom-paristech.fr).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2018.2889275

[1]SM is equal to the CM up to a scale factor when the latter exists.

These steps are repeated for a set of different parameters, which results in several estimates in each pixel. Finally, we perform bias-reduction and select the best estimate as in NL-SAR.

This letter is organized as follows. Section II introduces some background about PolSAR images and M-estimators. Section III presents the proposed method and the corresponding algorithm, with discussions. In Section IV, the results for simulated and RADARSAT-2 data are presented. Finally, some conclusions are drawn in Section V.

Notations: Vectors (resp. matrices) are denoted by bold-faced lowercase letters (resp. uppercase letters). ^{T} and ^{H}, respectively, represent the transpose and the Hermitian operator. $\mathrm{Tr}(\cdot)$ represents the trace of a matrix. Finally, $|\cdot|$ stands for the determinant of a matrix.

II. Theoretical Background

A. PolSAR

Polarimetric SAR sensors measure the amplitude and phase of backscattered signals in four combinations of the linear received and transmitted polarizations: horizontal–horizontal (HH), horizontal–vertical (HV), vertical–horizontal (VH), and vertical–vertical (VV). These signals form the complex scattering matrix $\mathbf{S}$

$$\mathbf{S} = \begin{bmatrix} S_{\mathrm{HH}} & S_{\mathrm{HV}} \\ S_{\mathrm{VH}} & S_{\mathrm{VV}} \end{bmatrix}$$

where S_{IJ} denotes the complex scattering amplitude for the received polarization I and the transmitted polarization J. The reciprocity theorem says that the cross-pol channels of the scattering matrix are equal, that is $S_{\mathrm{HV}} = S_{\mathrm{VH}}$. Therefore, there are only three independent complex coefficients required to characterize the SV $\mathbf{s} = [S_{\mathrm{HH}}\ \sqrt{2}S_{\mathrm{HV}}\ S_{\mathrm{VV}}]^T$ or alternatively, the SV is replaced by the linear transformation $\mathbf{k} = (1/\sqrt{2})[S_{\mathrm{HH}} + S_{\mathrm{VV}}\ S_{\mathrm{HH}} - S_{\mathrm{VV}}\ 2S_{\mathrm{HV}}]^T$ known as the Pauli representation of the SV [14].

B. PolSAR Modeling

So far, the most used model for the underlying distribution of PolSAR data is the complex Gaussian circular distribution. To reduce the effect of inherent speckle noise, PolSAR images are often spatially averaged, and the data are represented by the SCM $\widehat{\mathbf{\Sigma}}_{\mathrm{SCM}} = (1/N)\sum_{n=1}^{N} \mathbf{s}_n\mathbf{s}_n^H$ (or the sample coherency matrix $\widehat{\mathbf{\Sigma}}_{\mathrm{SCM}} = (1/N)\sum_{n=1}^{N} \mathbf{k}_n\mathbf{k}_n^H$). Under the assumption that $N > 3$, the estimate is complex Wishart distributed with N degrees of freedom (DoFs) around the expectation value $\mathbf{\Sigma}$ (true CM of Gaussian data).

Since the Gaussian distribution fails to give a good approximation in high-textured scenarios, an alternative is to use the SIRV model: $\mathbf{k} = \sqrt{\tau}\mathbf{n}$, where τ is a texture parameter whose distribution is unspecified, with mean value equal to 1, and $\mathbf{n}$ is the speckle vector, following a multivariate Gaussian distribution. SIRV represents a subclass of complex elliptically symmetric (CES) distributions. In this context, the performances of the SCM can be really degraded and one needs other estimators that are better adapted to a non-Gaussian framework.

C. M-Estimators

Let $(\mathbf{k}_1, \ldots, \mathbf{k}_N)$ be an N-sample of m-dimensional complex-independent CES-distributed vectors (with the same PDF). An M-estimator is defined by the solution of the following M-estimating equation:

$$\widehat{\mathbf{\Sigma}} = \frac{1}{N}\sum_{n=1}^{N} \varphi\left(\mathbf{k}_n^H \widehat{\mathbf{\Sigma}}^{-1}\mathbf{k}_n\right)\mathbf{k}_n\mathbf{k}_n^H \tag{1}$$

where φ is any real-valued weight function on $[0, \infty)$ [9] that does not need to be related to the PDF of the underlying distribution. The existence and uniqueness of the solution of (1) as well as the convergence of the corresponding recursive algorithm have already been shown, provided the function φ satisfies a set of general assumptions [11]. In particular, the resulting estimators are consistent estimators of the SM (up to a scale factor). Some of the widely used M-estimators are: Tyler's M-estimator [10], [15], Huber's M-estimator, and Student's M-estimator.

III. Robust M-NLMeans Method

In this section, we present step-by-step the proposed method for weights computation and discuss about its benefits.

A. Robust Preestimation

In order to compute the (dis)similarity between two pixels, one needs to compute a preestimation of the CM. In NL-SAR, this preestimation is done using truncated Gaussian on the patch of the size $S = (2s + 1) \times (2s + 1)$. In this method, we propose to use a Student's M-estimator given as the solution of

$$\widehat{\mathbf{\Sigma}}_t = \frac{m + \nu/2}{S}\sum_{n=1}^{S} \frac{\mathbf{k}_n\mathbf{k}_n^H}{\nu/2 + \mathbf{k}_n^H \widehat{\mathbf{\Sigma}}_t^{-1}\mathbf{k}_n} \tag{2}$$

where m is the vector dimension and ν is the DoF parameter. This estimator is very interesting, since it represents a sort of tradeoff between the SCM ($\nu \to \infty$) and Tyler's estimator ($\nu = 0$), the least and the most robust estimators. The SCM can be completely degraded with only one outlier while, on the other hand, Tyler's estimator is entirely resistant to aberrations but gives the estimation of so-called "shape" matrix, i.e., loses the information about matrix power. Therefore, Student's M-estimator represents a good compromise that simultaneously can be robust and preserve the power information.

B. Pixel Selection

Using these preestimated values, neighboring samples are selected around each pixel. The central pixel at location l is compared with all pixels in a circular window following a spiral path (see [1] for more details). To compute the dissimilarities between two pixels instead of classical generalized likelihood ratio tests, we propose to use the Box's M-test defined as

$$\mathcal{L} = \frac{|\widehat{\mathbf{\Sigma}}_1|^{S/2}|\widehat{\mathbf{\Sigma}}_2|^{S/2}}{|\widehat{\mathbf{\Sigma}}|^S}$$

where $\widehat{\mathbf{\Sigma}}_1$ is obtained with a sample $\mathbf{k}^{(1)} = (\mathbf{k}_1, \ldots, \mathbf{k}_S)$, $\widehat{\mathbf{\Sigma}}_2$ with $\mathbf{k}^{(2)} = (\mathbf{k}_{S+1}, \ldots, \mathbf{k}_{2S})$ and $\widehat{\mathbf{\Sigma}}$ with $\mathbf{k} = (\mathbf{k}^{(1)}, \mathbf{k}^{(2)})$.

This statistic has values between 0 and 1, where the values close to 0 reject the hypothesis that the matrices $\widehat{\mathbf{\Sigma}}_1$ and $\widehat{\mathbf{\Sigma}}_2$ are equal and values close to 1 accept it. By modifying the statistic $\mathcal{L}$, Box has obtained the approximated χ^2 distribution

$$u = -2(1-\beta)\ln(\mathcal{L}) \sim \chi^2(m(m+1)/2)$$

with $\beta = (3/2S)((2m^2+3m-1)/(6(m+1)))$, where m is the size of the SV, usually 3. We then propose to compute the similarity between two patches centered in pixels l and l' as

$$\Delta(l,l') = \sum_{\tau} u[(l+\tau),(l'+\tau)] \tag{3}$$

where $\tau \in [-p, p]$ is a 2-D shift indicating the location within each patch of size $P = (2p+1)\times(2p+1)$. Dissimilarities are then compared to the threshold in order to select similar pixels. Under the hypothesis $\mathcal{H}_0$ that two patches follow the same distribution, Δ has the χ^2 distribution with $d = (1/2)m(m+1)P$ DoF. The critical region of the test is then given by

$$\left\{R_c = \Delta, \Delta > \chi^2_{P_{fa}}(d)\right\}$$

with P_{fa} is the probability of false alarm and $\chi^2_{P_{\text{fa}}}(d)$ is the quantile of order $1-P_{\text{fa}}$ of $\chi^2(d)$.

C. Weights Computation

Once similar pixels are chosen, we proceed to the weights computation. To define the weights from the dissimilarity measure $\Delta(l,l')$, we propose to use an exponential kernel

$$\omega(l,l') = \begin{cases} e^{-\frac{|\Delta(l,l')-c|}{\lambda}} & \text{if } l \neq l' \\ 1 & \text{if } l = l'. \end{cases} \tag{4}$$

The parameter $c = \mathbb{E}[\Delta(l,l')|\mathcal{H}_0]$ is the expected dissimilarity of two patches under the $\mathcal{H}_0$ and the threshold λ can be computed as $\lambda = F^{-1}_{\chi^2(d)}(1-P_{FA})$. This mapping from the (dis)similarities to the weights prevents any pixel from having a larger weight than the central pixel. The parameter c has been introduced in order to give a weight close to 1 when the compared pixels come from the same distribution while preventing the noise enforcing. We normalize the quantity with λ in order to obtain comparable weight values for different values of p.

Finally, the weighted MLE is given by the weighted means

$$\widehat{\mathbf{\Sigma}}_{NL}(l) = \frac{\sum_{l'} \omega(l,l')\mathbf{k}'\mathbf{k}'^H}{\sum_{l'} \omega(l,l')}. \tag{5}$$

The method is recapped in Algorithm 1. First, the maximum sizes of search windows $\mathcal{W}$, patch $\mathcal{P}$, and preestimation scale $\mathcal{S}$ are set together with the threshold λ and the constant c that differs for each patch size. Then, the preestimation is performed for all pixels in the image and for all values of $0 \leq s \leq \mathcal{S}$, where $s = 0$ means the preestimation is off, i.e., the matrix is equal to $\mathbf{k}\mathbf{k}^H$ for the pixel at location l with coordinates (x, y). Then, for all window sizes, the central pixel is compared with all pixels in the window using the preestimations corresponding to different values of s and different sizes of patches to perform the patch comparison. Then, for each triple of (s, p, w), an estimate $\widehat{\mathbf{\Sigma}}_{\text{NL}}$ is computed. Afterward, for each $\widehat{\mathbf{\Sigma}}_{\text{NL}}$, the new one $\widehat{\mathbf{\Sigma}}_{\text{NLRB}}$ is obtained in order to reduce the bias introduced in the first one and finally, the best one (with the highest number of looks) is selected for each pixel giving the final filtered image. This final part, enclosed within the box, is the same as in NL-SAR and will not be detailed because of the lack of space.

Algorithm 1: *M*-NL Method

```
Initialization: W, P, S, λ, c, ν
forall the x, y do
    for s = 0 : S (scale size) do
        Preestimation with (2)
forall the x, y (coordinates of pixel l) do
    for w = 1 : W (search window size) do
        Compute Δx and Δy
        x' = x + Δx
        y' = y + Δy;                    ▷ coordinates of pixel l'
        for s = 0 : S do
            for p = 1 : P (patch size) do
                Compute Δ(l, l') with (3)
                if Δ(l, l') ≤ λ[p] then
                    Compute ω(l, l') with (4)
                else
                    ω(l, l') ← 0
    forall the s, p, w do
        Compute Σ̂_NL with (5)
        Bias-reduction step → Σ̂_NLRB
    return The best estimate
```

IV. Experiments: Implementation and Evaluation

A. Simulated Data

In this section, the results obtained for simulated and PolSAR data are presented. The simulated images have been generated using a Markov random field (MRF) following a Gibbs distribution as in [2]. Then, a polarimetric behavior has been assigned to the different parts of the designed images. The polarimetric signatures have been sampled from the PolSAR data as explained in the reference. The procedure is as follows. First, a random number C of polarimetric classes are chosen between 3 and 5, $C-1$ classes for distributed scatterers, and the last class for point scatters (targets). A ground truth is generated using an MRF and the targets correspond to squares of sizes varying between 2×2 and 5×5 pixels. For each distributed scatterer, one of the seven possible polarimetric signatures is randomly assigned and the Gaussian speckle noise is generated according to them. Finally, the targets generated using the remaining eighth polarimetric signature are added to the speckle noise. The exact values of CMs can be found in [2]. After the denoising, the set of following parameters has been evaluated.

1) *Radiometric Parameters σ:* Diagonal elements of the CM (power information).

2) *Complex Correlation Parameters ρ:* Derived from the three complex off-diagonal terms (channels correlation).
3) *Incoherent Decomposition Parameters:* Entropy (H), anisotropy (A), and the mean alpha angle ($\overline{\alpha}$)—scattering mechanism.
4) Co-polar and cross-polar polarization signatures (PSs) [16].
5) Edge preservation (EP).

For the first three groups of parameters, the estimated value is obtained as the corresponding mean value from the pixels for every scattering class (SC) and for every simulated image, given the filter. Then, the absolute relative bias of the estimated parameter is computed. The final (average) value is chosen as the median across all the simulated images and the various SCs. Then, the co- and cross-polar signatures of the average CM of the pixels belonging to a given SC of every simulated image are obtained (91×181 values). The average value for each class and image is obtained as a median value of absolute relative bias for every orientation and ellipticity angle. Then, as in the previous case, the final value is obtained as the median across all the images and classes. Finally, EP is measured on the boundary positions between extended targets. First, the gradient preservation (GP) is obtained as the average ratio of the observed gradient values on the power bands of the filtered power band image to the gradient values on the ground truth image. Then, the simple mapping from GP to EP is performed in order to give a measure close to 0 in the case of edge oversmoothing or undersmoothing and values close to 1 for good EP.

In order to perform the comparison, we have simulated one hundred 128×128 artificial PolSAR images as described previously. The set of parameters used in both methods is: window size: $\{3^2, 5^2, \ldots, 25^2\}$, patch size: $\{3^2, 5^2, \ldots, 11^2\}$, and scale: $\{0, 1, 2\}$. Since the speckle is Gaussian, we have chosen ν to be big enough ($\nu = 100$) in order to preserve the information about the texture and ensure the convergence of the solution in the preestimation step. The values for λ have been computed using the corresponding formula from Section III-C.

Fig. 1 shows the results of two different simulated images. The images are presented in the following order, from left to right and from top to bottom: original image (ground truth), speckle, NL-SAR results, and M-NL results. The first image [Fig. 1(a)] contains two classes for distributed scatterers. As can be seen from Fig. 1(c) and (d), the homogeneous areas are much better smoothed with M-NL than with NL-SAR while the edges are better captured and less blurred. The second image [Fig. 1(e)] consists of three SCs. In this case, we display the filtered images together with the difference of the filtered image and corresponding ground truth in order to better visualize the results. From Fig. 1(i) and (j), one can see that in some parts of the images, the noise is apparently more reduced with M-NL. This is the most visible (green points) at the image borders, and one can look, for instance, the top border and corners of the images. Obviously, the M-NL gives an estimation closer to the ground truth in both the cases. The visualization of the difference also reveals that most of the targets are better estimated with M-NL, which cannot be seen directly from Fig. 1(g) and (h). Some of them are marked in red squares.

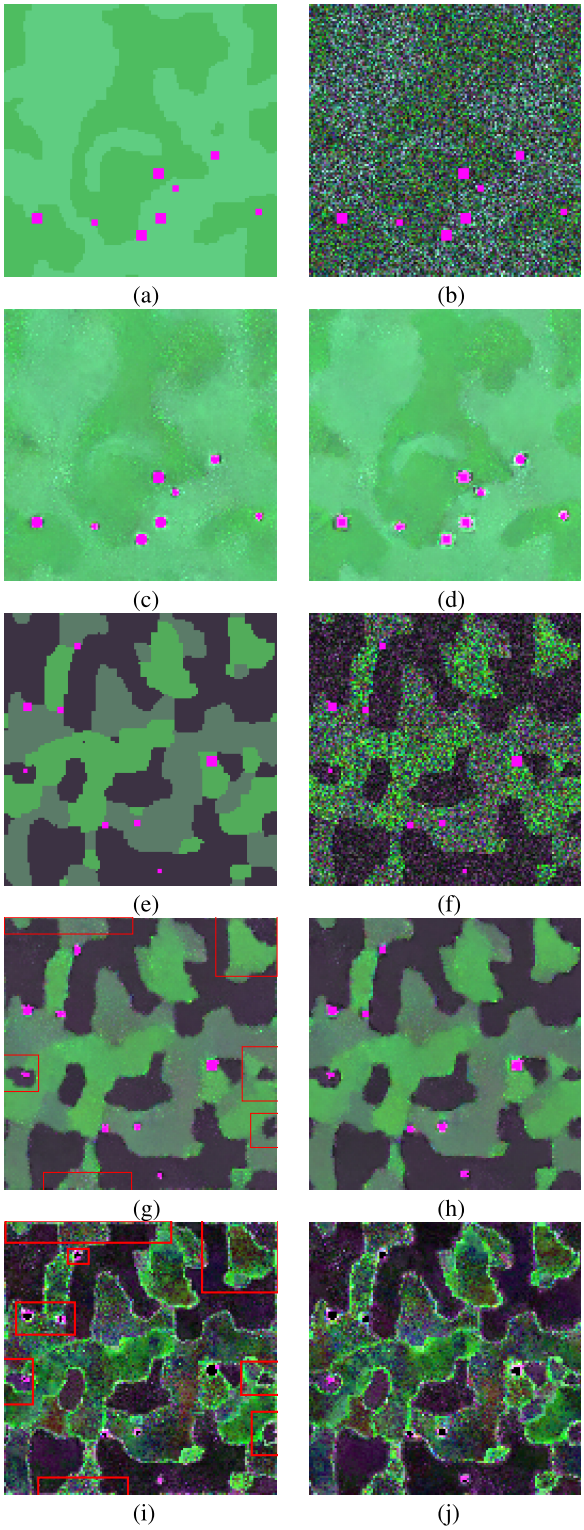

Fig. 1. Application to simulated data: results for two different realizations. (a) and (e) Ground truth. (b) and (f) Speckle. (c) and (g) Results obtained with NL-SAR. (d) and (h) Results obtained with M-NL. (i) and (j) Difference between the results and ground truth for NL-SAR and M-NL, respectively.

Table I lists the evaluation parameters defined above. The numerical results have been computed over the set of simulated PolSAR images and the final values are compared. One can

TABLE I

FILTERING RESULTS FOR SIMULATED DATA: ALL MEASURES BUT EP (EP $\in [0, 1]$) ARE ABSOLUTE RELATIVE ERRORS IN %

Filters	σ	$\lvert\rho\rvert$	$\angle\rho$	H
NL-SAR	2.21	**7.47**	**11.96**	14.51
M-NL	**1.56**	9.10	14.47	**14.49**
	$\overline{\alpha}$	A	PS	EP
NL-SAR	35.84	**10.51**	1.15	0.45
M-NL	**33.96**	10.97	**1.05**	**0.56**

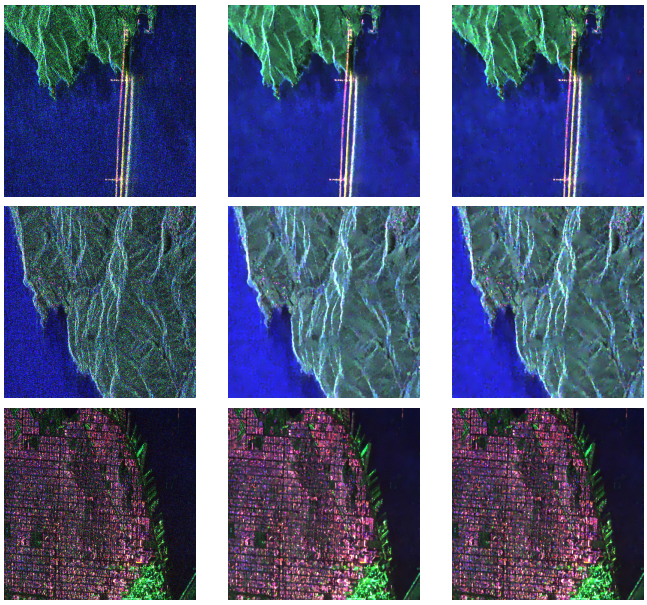

Fig. 2. Real data: San Francisco Bay—512 × 512 PolSAR images. (From left to right) Speckle, results obtained with NL-SAR, and results obtained with M-NL.

note that M-NL outperforms NL-SAR in almost all measures except ρ and A. Thus, it could be more convenient to use NL-SAR for terrain classification based on correlation coefficient or when measuring surface roughness implementing anisotropy parameter. On the other hand, M-NL gives a better estimation of radiometric parameters, almost all incoherent decomposition parameters, PSs, and EP parameters. A significant improvement in EP is also visually noticeable on the simulated images, and thus these results are not surprising.

B. RADARSAT-2 PolSAR Data

Finally, the results for real data are given. Three different parts of San Francisco Bay are presented from top to bottom representing different scenarios in PolSAR images, such as water, vegetation, and urban areas. In this case, we do not dispose any information about the ground truth, and thus one can analyze the results only visually. First, one can note that M-NL better smoothes the homogeneous areas, while preserving well the edges in textured scenarios. It can also be noted that, as in the case of simulated data, M-NL gives results with higher contrast.

Last but not least, we would like to mention that an advantage of this method is also its simplicity, since, after the preestimation, it immediately starts to measure the dissimilarities and computes weights using a single, simple formula. On the other side, NL-SAR first needs to analyze noise in a homogeneous area in order to learn its distribution. Then, when the denoising step starts, a binary search is necessary to find the corresponding quantile and then a mapping is used to evaluate the weight, which obviously requires more steps.

V. CONCLUSION

This letter has investigated a new statistical approach for NL estimation in PolSAR imagery. The proposed method relies on M-estimators and has been compared with the NL-SAR method showing better results. It should be noted that the proposed method has common steps with NL-SAR. Yet, the estimation of NL estimates, the crucial part of any NL denoising method, is based on the statistical behavior of M-estimators, which improves the method simplicity.

REFERENCES

[1] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger, "NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In)SAR denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2021–2038, Apr. 2015.

[2] S. Foucher and C. López-Martínez, "Analysis, evaluation, and comparison of polarimetric SAR speckle filtering techniques," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1751–1764, Apr. 2014.

[3] J.-S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, Sep. 1999.

[4] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[5] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2. Washington, DC, USA: IEEE Computer Society, Jun. 2005, pp. 60–65.

[6] C.-A. Deledalle, F. Tupin, and L. Denis, "Polarimetric SAR estimation based on non-local means," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 2515–2518.

[7] J. Chen, Y. Chen, W. An, Y. Cui, and J. Yang, "Nonlocal filtering for polarimetric SAR data: A pretest approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1744–1754, May 2011.

[8] G. Vasile, J. P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1809–1826, Apr. 2010.

[9] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *Ann. Statist.*, vol. 4, no. 1, pp. 51–67, 1976.

[10] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *Ann. Statist.*, vol. 15, no. 1, pp. 234–251, 1987.

[11] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5597–5625, Nov. 2012.

[12] G. Drašković and F. Pascal, "New properties for Tyler's covariance matrix estimator," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2016, pp. 820–824.

[13] G. Drašković and F. Pascal, "New insights into the statistical properties of M-estimators," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4253–4263, Apr. 2018.

[14] J. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL, USA: CRC Press, 2009.

[15] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound-Gaussian noise: Existence and algorithm analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 34–48, Jan. 2008.

[16] J. J. van Zyl, H. A. Zebker, and C. Elachi, "Imaging radar polarization signatures: Theory and observation," *Radio Sci.*, vol. 22, no. 4, pp. 529–543, Aug. 1987.