Joint despeckling and thermal noise compensation: application to Sentinel-1 images of the Arctic

Inès Meraoumia, Debanshu Ratha, Emanuele Dalsasso, Johannes Lohse, Florence Tupin, Andrea Marinoni, Loïc Denis

Joint despeckling and thermal noise compensation: application to Sentinel-1 images of the Arctic

Inès Meraoumia, Debanshu Ratha, Emanuele Dalsasso, Johannes Lohse,
Florence Tupin, Andrea Marinoni, Loïc Denis

***Abstract*—Synthetic Aperture Radar (SAR) images offer crucial information for studying and monitoring sea ice in the Arctic. Sentinel-1 captures images of the area using an extremely wide swath for reduced revisit time. The backscattered signal from sea ice and open water is often very weak, making it difficult to distinguish from the sensor thermal noise floor. Thermal noise impacts the images by generating a bias and increasing the fluctuations related to speckle phenomenon. Analyzing these images requires both correcting this bias and reducing fluctuations without blurring out the image content. The acquisition of several sub-swaths in a single pass using Terrain Observation with Progressive Scans (TOPS) produces images that exhibit, after compensation for antenna gains, a non-uniform thermal noise floor and strong discontinuities between sub-swaths. Denoising techniques must take these specificities into account to restore the images.**

This paper introduces a joint approach to remove the thermal noise offset and suppress fluctuations due to speckle and thermal noise. Compensating at once for all these effects largely reduces artifacts at the boundary between sub-swaths. We demonstrate using both numerical simulations and actual Sentinel-1 images that debiased polarimetric reflectivities can be recovered and fluctuations strongly reduced while preserving fine spatial structures.

***Index Terms*—SAR, Despeckling, Thermal noise, Sea Ice**

I. Introduction

Thanks to its ability to see through clouds and operate regardless of daylight conditions, SAR imaging plays a key role in Earth observation. With its side-looking geometry, it can cover swaths several hundred kilometers wide. The diversity of available spatio-temporal system configurations [1] and the various polarimetric and interferometric capabilities [2], [3] open a very wide scope of applications: from the study of forests, agriculture, cryosphere, and urban areas to oceans [4].

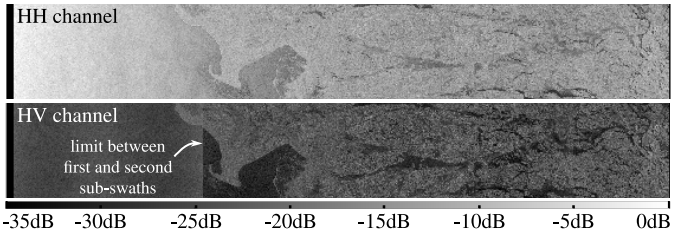


Figure 1. A Sentinel-1 EW GRDM image with both open water (on the left) and sea ice (on the right). Thermal noise dominates the water backscattering in the HV cross-polarization channel, especially in the first sub-swath.

Studies report that the Arctic is warming up at a quadrupled pace compared to the rest of the globe since 1979 [5]. The study of sea ice concentration, its extent, ice type, the melting of northern glaciers and the calving of the Greenland glacier ice sheet are all of crucial interest for the study of climate change but also for maritime navigation in the area. SAR imaging is very well suited to monitor the Arctic region: longer dark periods, frequent precipitation, and cloud cover highly disadvantage the acquisition by optical sensors [6]–[8].

The Sentinel-1 constellation from the European Space Agency (ESA) uses a special Extra Wide (EW) swath mode to cover the Arctic area with a short revisit time. Each image covers a 400 km wide swath made of five sub-swaths that are swept sequentially in short bursts thanks to digital antenna steering and the Terrain Observation with Progressive Scans (TOPS) principle [9], [10]. Operational sea-ice classification methods using dual-polarization Sentinel-1 EW data are still relatively scarce [11]. Analyzing these data is indeed challenging for at least two reasons: (i) there is a high dependence of Sentinel-1 C-band backscatter to the incidence angle [12], which needs to be included by the sea ice classification method [13], and (ii) the noise is very strong [11]. Beyond the speckle phenomenon that corrupts all SAR images, the proximity of sea ice backscattered signal to the thermal noise floor of the sensor indeed strongly affects the data. Improving the signal-to-noise ratio in these images requires the removal of the bias due to the thermal noise floor and the suppression of fluctuations originating both from the speckle phenomenon and from thermal noise. The issue of correcting the bias and reducing the fluctuations are often addressed separately in the literature.

To compensate for the two-way antenna gain, measurements collected in directions away from the main diffraction lobe of

I. Meraoumia and F. Tupin are with LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France, (e-mail: forename.name@telecom-paris.fr).

J. Lohse was with the Department of Physics and Technology, UiT The Arctic University of Norway, 9019 Tromsø, Norway, and is currently at the Australian Antarctic Program Partnership, Institute for Marine and Antarctic Studies, University of Tasmania, Nipaluna / Hobart, Tasmania (johannes.p.lohse@uit.no).

D. Ratha is with Akvaplan-niva AS and formerly worked at IET and IFT, UiT The Arctic University of Norway. (email: der@akvaplan.niva.no).

E. Dalsasso is with Environmental Computational Science and Earth Observation (ECEO) laboratory, EPFL, 1951 Sion Switzerland. (e-mail: emanuele.dalsasso@epfl.ch).

A. Marinoni was with the Department of Physics and Technology, UiT The Arctic University of Norway, 9019 Tromsø, Norway, and is currently with the Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, UK (email: am2920@cam.ac.uk).

L. Denis is with Université Jean Monnet Saint-Etienne, CNRS, Institut d Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France and also with LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France, (e-mail: loic.denis@univ-st-etienne.fr).

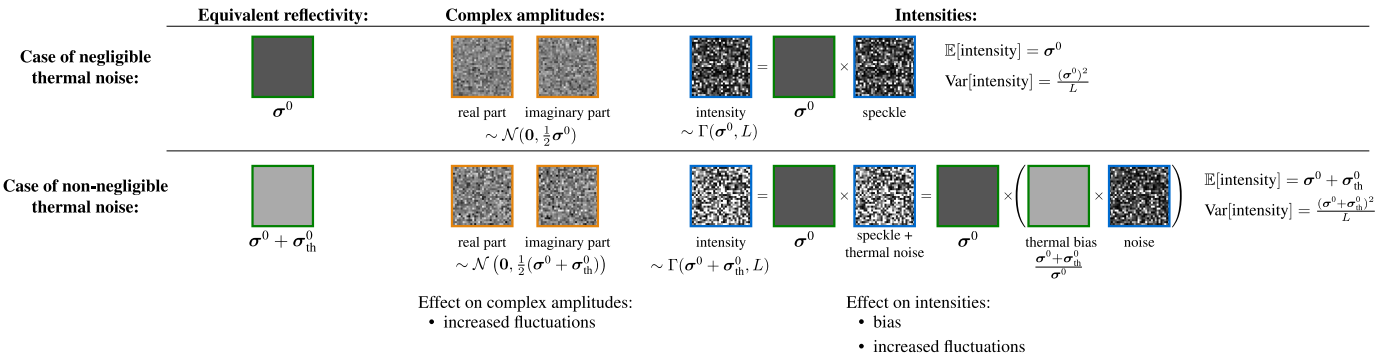


Figure 2. Illustration of the effect of thermal noise on radar intensities: when thermal noise is negligible, the intensities fluctuate according to a gamma distribution, the expectation matches the reflectivity $\boldsymbol{\sigma}^0$ and the variance is proportional to $(\boldsymbol{\sigma}^0)^2$; when thermal noise is no longer negligible, a factor $(\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^0_{\text{th}})/\boldsymbol{\sigma}^0$ appears, producing a bias (the thermal noise level $\boldsymbol{\sigma}^0_{\text{th}}$ adds up to the scene reflectivity $\boldsymbol{\sigma}^0$) and amplifying the fluctuations (the variance is multiplied by $(\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^0_{\text{th}})^2/(\boldsymbol{\sigma}^0)^2$, which is larger than one).

the antenna are numerically amplified. This turns the constant thermal noise floor into a spatial pattern that depends on the range and azimuth coordinates. Particularly noticeable is the discontinuity between the first and second sub-swaths, see Figure 1. The bias due to thermal noise impacts sigma naught values most significantly over low backscatter zones (e.g. leads/open water, young smooth ice) [14]. Removing this bias requires an estimation of the thermal noise floor pattern, also known as the Noise Equivalent Sigma Zero (NESZ). For sea ice and oceanographic applications, the calibrated noise vectors provided with the ground range detected (GRD) product data by ESA have not been found fully satisfactory and several approaches to improve this correction have been developed. Several of these methods rely on detecting open water areas where the thermal noise dominates and fitting a parametric model of the NESZ map [15], [16]. Some works also compensate for the fluctuations due to the thermal noise by rescaling the image [17]–[19]. A limitation of the fitting approaches is that the reflectivity of the radar scene degrades the estimation, making the correction imperfect for areas covered by sea ice. In our experiments, we found that the method of Korosov et al. in [20], based on an improvement of [16], led to good estimates of the NESZ map and focus in this paper on combining this approach with despeckling techniques.

To minimize the impact of fluctuations caused by both the speckle phenomenon and the thermal noise in SAR images, the use of despeckling techniques is necessary. Despeckling is a long-standing topic in radar imaging and numerous approaches have been proposed [21]–[24]. The state-of-the-art uses deep neural networks. Due to the difficulty of obtaining ground truth speckle-free SAR images, self-supervised learning has appeared as a relevant paradigm for developing despeckling algorithms [25], [26]. The key idea of self-supervised learning is the splitting of noisy observations in two subsets related to the same radar scene but where the speckle component is statistically independent. It is then possible to train a neural network by confronting predictions made using only information from the first subset with the noisy observations in the second subset. Data splitting can be performed by separating pairs of images of the same area captured at different dates [27], [28], excluding a pixel from its spatial neighborhood [29], or decomposing a single-look complex image into its real and imaginary components [30].

Existing methods either tackle the problem of refining the NESZ map to improve the debiasing step or consider only the despeckling problem. We show in this paper that the straightforward combination of these two operations in two steps leads to difficulties: on the one hand, subtracting the NESZ map changes the statistics of speckle so that existing despeckling methods become inapplicable, on the other hand, despeckling the image before removing the NESZ offset leads to artifacts because discontinuities in the NESZ map are not perfectly recovered by the despeckling algorithm.

Our contributions: We develop a method to remove both the thermal noise bias and the fluctuations due to speckle and thermal noise. Our method offers the following distinctive features:

1) it is derived from a generative model, introduced in Section II-A, that accounts for the physics of Sentinel-1 dual-polarization imaging,
2) the despeckling step, described in Section II-C, is aware of the NESZ map so that it can prevent artifacts at discontinuities between sub-swaths and account for the increased variance due to thermal noise in areas of low reflectivity,
3) it processes jointly the two polarimetric channels to exploit their redundancy and better restore fine details and low-contrast structures,
4) we develop a self-supervised learning scheme that only requires co-registered ground-detected images for training and a single image at inference.

II. Proposed method: learning to suppress speckle and thermal noise contributions

This section derives a generative model to account for both the speckle phenomenon and the thermal noise. This model is key not only to analyze how images can be restored (Section II-B), but also to generate images with synthetic noise to perform the first training step of our restoration network, described in Section II-C, and, in Section III-A, to conduct a

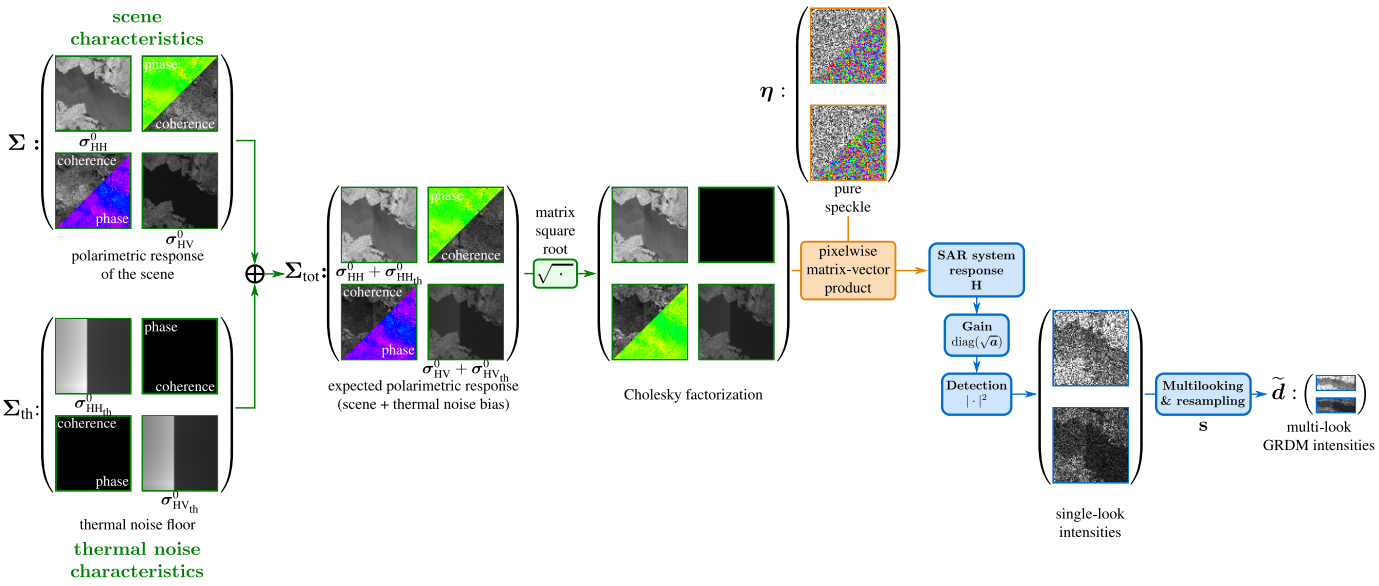


Figure 3. A generative model statistically equivalent to the physics of Sentinel-1 dual-pol GRD imagery: the effects of speckle and thermal noise are modeled through an equivalent covariance matrix $\mathbf{\Sigma}_{\text{tot}}$.

quantitative evaluation of restoration methods with a ground-truth image.

A. *Model of the thermal noise component in Sentinel-1 images*

Thermal noise sets a limit on the sensitivity of a SAR system: the in-phase and quadrature measurements are corrupted by an additive white Gaussian noise, and radar echoes with an amplitude much smaller than the noise floor are then almost impossible to recover. Since the speckle phenomenon can also be modeled as complex circular Gaussian fluctuations, thermal noise appears, in synthesized SAR images, as a background component with a low reflectivity $\boldsymbol{\sigma}^0_{\text{th}}$ that adds to the reflectivity $\boldsymbol{\sigma}^0$ of the SAR scene. Fluctuations of the intensity in the SAR images are then proportional to $\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^0_{\text{th}}$, see Figure 2. While areas with strong reflectivities (where $\boldsymbol{\sigma}^0 \gg \boldsymbol{\sigma}^0_{\text{th}}$) are not significantly affected by thermal noise, lower-reflectivity regions suffer from increased fluctuations compared to the level of fluctuations that would occur due to speckle alone. Furthermore, estimations of the reflectivity are biased unless the thermal floor level is removed.

When acquiring SAR images over a large swath, it is necessary to compensate for the non-uniform two-way antenna gain over the swath. After this correction, the noise equivalent sigma zero $\boldsymbol{\sigma}^0_{\text{th}}$ is no longer constant. In images captured using several sub-swaths (ScanSAR and TOPS techniques [9]), values of $\boldsymbol{\sigma}^0_{\text{th}}$ display strong discontinuities at the boundary between two sub-swaths or between bursts (see Figure 1). Removing the thermal noise bias $\boldsymbol{\sigma}^0_{\text{th}}$ is necessary to remove those discontinuities that are particularly visible in areas with low reflectivity.

Figure 3 gives an approximate generative model for polarimetric SAR imaging[1] that includes the two main perturbation components in SAR images: thermal noise and speckle. Many works have been devoted to the statistical modeling of interferometric or polarimetric covariance matrices, see for example [31]–[34]. The model derived for multi-temporal stacks in [35] is extended here to account for thermal noise, the antenna pattern, detection, ground projection, multilooking and resampling. It will be used in the following to produce synthetically degraded images for the first training step of our restoration network. It also provides useful insight into the respective impact of thermal noise and speckle. The left-hand side of the figure illustrates the components of the polarimetric covariance matrix that characterizes the scene $\mathbf{\Sigma}$ and the thermal noise floor $\mathbf{\Sigma}_{\text{th}}$. These two independent components add up to form the total covariance matrix $\mathbf{\Sigma}_{\text{tot}}$. In the case of dual-pol imaging, matrices $\mathbf{\Sigma}$, $\mathbf{\Sigma}_{\text{th}}$, and $\mathbf{\Sigma}_{\text{tot}}$ are of size $2N \times 2N$ for N-pixels images. Since thermal noise on the HH and HV measurements is independent, off-diagonal values of $\mathbf{\Sigma}_{\text{th}}$ are null. While pure speckle follows a complex circular Gaussian distribution $\mathcal{N}_c(\mathbf{I})$ with a unitary covariance matrix $\mathbf{I}$, single-look complex polarimetric SAR images follow the distribution $\mathcal{N}_c(\mathbf{\Sigma}_{\text{tot}})$. If $\mathbf{M}_{\text{tot}}$ is a matrix such that $\mathbf{M}_{\text{tot}}\mathbf{M}^*_{\text{tot}} = \mathbf{\Sigma}_{\text{tot}}$, then pure speckle $\boldsymbol{\eta} \sim \mathcal{N}_c(\mathbf{I})$ can be turned into complex amplitudes $\boldsymbol{z} = \mathbf{M}_{\text{tot}}\boldsymbol{\eta}$ with $\boldsymbol{z} \sim \mathcal{N}_c(\mathbf{\Sigma}_{\text{tot}})$, i.e., the factorization of matrix $\mathbf{\Sigma}_{\text{tot}}$ provides a generative model of complex-valued polarimetric SAR amplitudes. In Figure 3 we illustrate the Cholesky factorization[2] [36], defined at a given pixel n by:

$$\begin{aligned}\mathbf{\Sigma}_{\text{tot}_n} &= \begin{pmatrix} \sigma^0_{\text{HH}_n} + \sigma^0_{\text{HH}_{\text{th}\,n}} & \sqrt{\sigma^0_{\text{HH}_n}\sigma^0_{\text{HV}_n}}\rho_n e^{j\beta_n} \\ \sqrt{\sigma^0_{\text{HH}_n}\sigma^0_{\text{HV}_n}}\rho_n e^{-j\beta_n} & \sigma^0_{\text{HV}_n} + \sigma^0_{\text{HV}_{\text{th}\,n}} \end{pmatrix} \\ &= \mathbf{M}_{\text{tot}_n}\mathbf{M}^*_{\text{tot}_n}\,, \end{aligned} \tag{1}$$

[1] since our method is illustrated on Sentinel-1 data, the simplified case of dual-polarimetry is shown here, the extension of the model to full-polarimetry is discussed in Appendix A

[2] other factorizations are also possible such as $\mathbf{M}_{\text{tot}} = (\mathbf{\Sigma}_{\text{tot}})^{1/2}$; the Cholesky factorization given in Eq. (2) can be easily checked by forming the product $\mathbf{M}_{\text{tot}}\mathbf{M}^*_{\text{tot}}$.

where $\rho_n \in [0,1[$ and $\beta_n \in [-\pi,\pi[$ are the polarimetric coherence and polarimetric phase at pixel n, and with

$$\mathbf{M}_{\text{tot}_n} = \begin{pmatrix} \sqrt{\sigma^0_{\text{HH}_n} + \sigma^0_{\text{HH}_{\text{th}\,n}}} & 0 \\ \sqrt{u_n}\exp(-j\beta_n) & \sqrt{\sigma^0_{\text{HV}_n} + \sigma^0_{\text{HV}_{\text{th}\,n}} - u_n} \end{pmatrix} \tag{2}$$

where

$$u_n = \frac{\sigma^0_{\text{HH}_n}\sigma^0_{\text{HV}_n}}{\sigma^0_{\text{HH}_n} + \sigma^0_{\text{HH}_{\text{th}\,n}}}\rho_n^2\,. \tag{3}$$

The full matrices $\mathbf{\Sigma}_{\text{tot}}$ and $\mathbf{M}_{\text{tot}}$ are block diagonal, with blocks $\mathbf{\Sigma}_{\text{tot}_n}$ and $\mathbf{M}_{\text{tot}_n}$: polarimetric channels are correlated but pixels are spatially independent.

The expression of this factorization will be useful to simulate synthetic speckle in the following paragraphs, both as a first step to perform the initial training of our deep neural network (Section II-C) and to perform quantitative evaluation of different methods in a controlled setting with known ground truth (Section III-A).

The right-hand-side of Figure 3 illustrates the steps that lead to ground-detected images such as Sentinel-1 GRDM data. Starting from the complex amplitudes of an ideal synthesized SAR image, the SAR system response $\mathbf{H}$ (modeled as a linear operator in the spatial domain) introduces spatial correlations (due to possible zero-padding and spectral apodization during the SAR synthesis), then a gain $\sqrt{\boldsymbol{a}}$ transforms amplitudes into digital numbers, and the detection step suppresses the phase information. Finally, a multilooking and resampling step is performed in order to obtain approximately square pixels and images with a smaller memory footprint (linear filtering operation $\mathbf{S}$). In equation, the generative model of Figure 3 writes:

$$\widetilde{\boldsymbol{d}} = \mathbf{S}|\text{diag}(\sqrt{\boldsymbol{a}})\mathbf{H}\boldsymbol{z}|^2\,, \tag{4}$$

where $\widetilde{\boldsymbol{d}}$ is a vector containing the GRD in medium resolution (GRDM) intensities (i.e., the input data in our image restoration problem), $\mathbf{S}$ is an operator that multilooks then subsamples to transform single-look intensities into GRDM images, the squared modulus is applied separately to each complex value, and $\boldsymbol{z}$ is the vector formed by the concatenation of the polarimetric complex amplitudes $\boldsymbol{z}_n$ at each pixel of the image. These polarimetric complex amplitudes $\boldsymbol{z}_n$ are formed by the pixel-wise operations

$$\forall n \in [\![1,N]\!],\ \boldsymbol{z}_n = \mathbf{M}_{\text{tot}_n}\boldsymbol{\eta}_n\,, \tag{5}$$

where the pixel number n appears as an index to indicate the restriction to a single pixel location, and the matrix $\mathbf{M}_{\text{tot}_n}$ is formed by the matrix factorization

$$\mathbf{M}_{\text{tot}_n}\mathbf{M}^*_{\text{tot}_n} = \mathbf{\Sigma}_{\text{tot}_n} = \mathbf{\Sigma}_{n} + \mathbf{\Sigma}_{\text{th}_n}\,. \tag{6}$$

At the spatial resolution and pixel spacing of GRDM images, denoted with $\widetilde{\cdot}$ on the variables, the bias due to the thermal noise floor can be removed, leading to calibrated and corrected data $\widetilde{\boldsymbol{d}}_{\text{c}}$ [20]:

$$\widetilde{\boldsymbol{d}}_{\text{c}} = \text{diag}(1/\widetilde{\boldsymbol{a}})\widetilde{\boldsymbol{d}} - \widetilde{\boldsymbol{\sigma}}^0_{\text{th}}\,, \tag{7}$$

where the map of the inverse gain $1/\widetilde{\boldsymbol{a}}$ is obtained by pixelwise division and the gain $\widetilde{\boldsymbol{a}}$ and noise equivalent sigma zero $\widetilde{\boldsymbol{\sigma}}^0_{\text{th}}$ at the GRDM resolution are obtained by application of the resampling operator $\mathbf{S}$: $\widetilde{\boldsymbol{a}} = \mathbf{S}\boldsymbol{a}$ and $\widetilde{\boldsymbol{\sigma}}^0_{\text{th}} = \mathbf{S}\boldsymbol{\sigma}^0_{\text{th}}$, and typically made available by the space agencies in the metadata that accompany the SAR images.

The aim of the despeckling and thermal noise compensation is to remove both the thermal bias and the fluctuations due to thermal noise and speckle from the data $\widetilde{\boldsymbol{d}}$, i.e., to build estimators of the expectation $\mathbb{E}[\widetilde{\boldsymbol{d}}_{\text{c}}]$. As shown below, this expectation corresponds to the reflectivity of the scene, up to the low-pass filtering effect of the SAR system. Combining equations (7) and (4) gives:

$$\mathbb{E}[\widetilde{\boldsymbol{d}}_{\text{c}}] = \text{diag}(1/\widetilde{\boldsymbol{a}})\mathbb{E}\left[\mathbf{S}|\text{diag}(\sqrt{\boldsymbol{a}})\mathbf{H}\boldsymbol{z}|^2\right] - \widetilde{\boldsymbol{\sigma}}^0_{\text{th}} \tag{8}$$

Since calibration factors $\boldsymbol{a}$ vary slowly with the range and azimuth location, $\mathbf{S}|\text{diag}(\sqrt{\boldsymbol{a}})\mathbf{H}\boldsymbol{z}|^2 \approx \text{diag}(\widetilde{\boldsymbol{a}})\mathbf{S}|\mathbf{H}\boldsymbol{z}|^2$ and the two terms $\text{diag}(1/\widetilde{\boldsymbol{a}})$ and $\text{diag}(\widetilde{\boldsymbol{a}})$ cancel out. The expectation $\mathbb{E}\left[|\mathbf{H}\boldsymbol{z}|^2\right]$ can be rewritten $\text{diag}\left(\mathbf{H}\,\mathbb{E}\left[\boldsymbol{z}\boldsymbol{z}^*\right]\mathbf{H}^*\right)$, where the notation diag() refers to the extraction of the diagonal of a square matrix in that context. The linear operator $\mathbf{H}$ operates separately on each polarimetric channel. Moreover, complex amplitudes in $\boldsymbol{z}$ are uncorrelated for any pair of distinct pixels. At pixel n, the expectation is thus equal to $\sum_k |\mathrm{H}_{nk}|^2 \cdot (\sigma^0_{\text{HH}_k} + \sigma^0_{\text{HH}_{\text{th}\,k}})$, i.e., the low-pass filtered reflectivity obtained by accounting for the incoherent point spread function of the SAR system. Provided that $\widetilde{\boldsymbol{\sigma}}^0_{\text{th}}$ matches the resampled and low-pass filtered noise equivalent sigma zero values in the HH and HV polarization channels ($\widetilde{\boldsymbol{\sigma}}^0_{\text{th}} = \mathbf{S}|\mathbf{H}|^2\boldsymbol{\sigma}^0_{\text{th}}$), we obtain:

$$\mathbb{E}[\widetilde{\boldsymbol{d}}_{\text{c}}] \approx \mathbf{S}|\mathbf{H}|^2\boldsymbol{\sigma}^0\,, \tag{9}$$

where the linear operator[3] $|\mathbf{H}|^2$ can be interpreted, for a shift-invariant SAR imaging system (e.g., operating in stripmap mode), as a convolution by the squared modulus of the complex-valued impulse response (a product of cardinal sines, in the absence of spectral apodization). To conclude, our objective is to recover the scene $\widetilde{\boldsymbol{\sigma}}^0$, in ground range geometry, from the ground detected data $\widetilde{\boldsymbol{d}}$ by removing thermal bias and fluctuations due both to thermal noise and speckle. This scene is a low-pass filtered and resampled version of the scene in slant geometry $\boldsymbol{\sigma}^0$: $\widetilde{\boldsymbol{\sigma}}^0 = \mathbf{S}|\mathbf{H}|^2\boldsymbol{\sigma}^0$ with $\mathbf{S}|\mathbf{H}|^2$ a linear operator that applies the incoherent point spread function of the system, low-pass filters and resamples.

The effect of thermal noise is mostly seen in areas with low reflectivities, such as water and ice areas in Figure 1. To quantify this effect, we can compare the quadratic errors $\epsilon^2_{\text{HH}n} = \mathbb{E}[(\widetilde{\sigma}^0_{\text{HH}_n} - \widetilde{d}_{\text{HH}_n}/\widetilde{a}_{\text{HH}n})^2]$ and $\epsilon^2_{\text{HV}n} = \mathbb{E}[(\widetilde{\sigma}^0_{\text{HV}_n} - \widetilde{d}_{\text{HV}_n}/\widetilde{a}_{\text{HV}n})^2]$ in the presence or in the absence of thermal noise. The bias of $\widetilde{d}_{\text{HH}_n}/\widetilde{a}_{\text{HH}n}$ corresponds to $\widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}$ and the variance to $\frac{1}{L}\left(\widetilde{\sigma}^0_{\text{HH}n} + \widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}\right)^2$. By noting $\text{SNR}_{\text{HH}n}$ the signal-

[3] the squared modulus is applied elementwise to matrix $\mathbf{H}$, leading to another matrix of same dimension

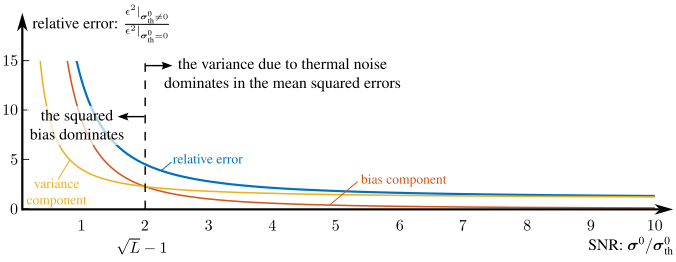


Figure 4. Impact of the thermal noise on the radar intensities: evolution of the relative mean squared errors ϵ^2 with the signal-to-noise ratio, as established in equations (10) and (11).

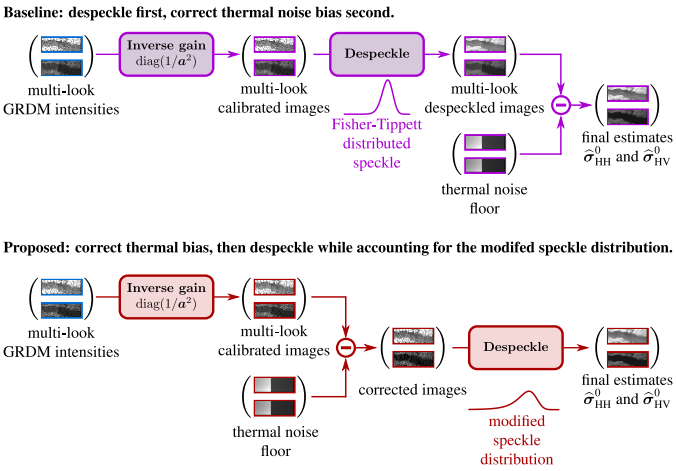


Figure 5. Despeckling techniques generally assume a multiplicative speckle model and must be applied *before* thermal noise bias removal (top). We propose to remove this bias before the despeckling step, provided that a specific despeckling technique be applied (bottom). This improves the restoration because the thermal noise floor discontinuities are removed before the despeckling step, preventing the apparition of artifacts.

to-noise ratio $\widetilde{\sigma}^0_{\text{HH}_n}/\widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}$, we obtain:

$$\frac{\epsilon^2_{\text{HH}_n}\big|_{\widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}\neq 0}}{\epsilon^2_{\text{HH}_n}\big|_{\widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}=0}} = \frac{\left(\widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}\right)^2 + \frac{1}{L}\left(\widetilde{\sigma}^0_{\text{HH}_n} + \widetilde{\sigma}^0_{\text{HH}_{\text{th}\,n}}\right)^2}{\frac{1}{L}\left(\widetilde{\sigma}^0_{\text{HH}_n}\right)^2} = \frac{L}{\text{SNR}^2_{\text{HH}_n}} + \left(1+\frac{1}{\text{SNR}_{\text{HH}_n}}\right)^2 \tag{10}$$

and

$$\frac{\epsilon^2_{\text{HV}_n}\big|_{\widetilde{\sigma}^0_{\text{HV}_{\text{th}\,n}}\neq 0}}{\epsilon^2_{\text{HV}_n}\big|_{\widetilde{\sigma}^0_{\text{HV}_{\text{th}\,n}}=0}} = \frac{L}{\text{SNR}^2_{\text{HV}_n}} + \left(1+\frac{1}{\text{SNR}_{\text{HV}_n}}\right)^2. \tag{11}$$

These expressions, plotted in Figure 4, indicate that when the number of looks L becomes large, the dominant impact of thermal noise is the bias. For a fixed value of L, thermal noise leads to an excess of error due to the combination of bias and an increased variance. When SNR $< \sqrt{L}-1$ the bias is the most impacting factor, when SNR $> \sqrt{L}-1$ the variance increase dominates. To properly reduce the errors due to thermal noise, it is, therefore, necessary to address both issues, which motivates the design of a method that performs jointly the despeckling and thermal noise compensation.

B. Proposed joint despeckling and thermal noise bias correction

The most straightforward way to reduce speckle and thermal noise fluctuations and compensate as well for the bias due to thermal noise is a sequential processing that reverts the steps of the generative model depicted in Figure 3: after calibrating the image, first despeckling, leading to over-estimated reflectivities, then subtracting the thermal noise floor $\boldsymbol{\sigma}^0_{\text{th}}$. This strategy is illustrated in the top row of Figure 5 and constitutes our baseline. The speckle filter used in the baseline in based on the SAR2SAR framework [27], trained specifically for dual-pol Sentinel-1 images in EW mode [37]. The main weakness of this strategy is that discontinuities of $\boldsymbol{\sigma}^0_{\text{th}}$ are present in the image processed by the despeckling algorithm. Imperfect restoration of these edges leads to artifacts after the subtraction step (as shown in Section III). A better approach would consist in removing the thermal noise bias *before* performing the reduction of fluctuations due to speckle and thermal noise. This, however, changes dramatically the statistics of speckle in the areas with low SNR: Figure 6 illustrates how the distribution of log-intensities, modeled by Fisher-Tippett distribution (red dashed curve) gets distorted once the thermal noise floor has been removed. Conventional despeckling techniques can no longer be applied to these corrected images. Yet, training a dedicated despeckling network is possible: this is the approach illustrated at the bottom of Figure 5 that we propose in this paper.

C. Self-supervised training with image pairs: the SAR2SAR framework

In section I, three self-supervised training strategies were described: (i) the use of pairs of images of the same area, captured at different times so that speckle realizations can be considered independent [27]; (ii) the use of all neighboring values except the central pixel, using a network with an architecture leading to a receptive field with a blind spot [29]; (iii) the decomposition of single-look complex images into real and imaginary components [30].

The approach (iii) cannot be directly applied to Sentinel-1 GRDM images since only intensities are available. Following this approach would require going back to the single-look complex images and repeating the pipeline that produces medium-resolution ground-detected images, a process that may be too heavy for operational use. Due to the spatial averaging steps corresponding to the linear operators $\mathbf{H}$ and $\mathbf{S}$ in the model (4), the approach (ii) would not work: speckle is spatially correlated in Sentinel-1 GRDM images, which breaks the independence assumption between neighboring pixels at the core of the method. Only approach (i) lends itself to the training of a network for Sentinel-1 GRDM images.

Thanks to the revisit of the satellite, it is possible to collect several images of the same area. If the temporal separation between the images is sufficient, speckle almost completely decorrelates and the images can be considered as independent random draws. Yet, scene evolution over time necessitates consideration of these changes before comparing the despeckled estimate derived from the first image to the

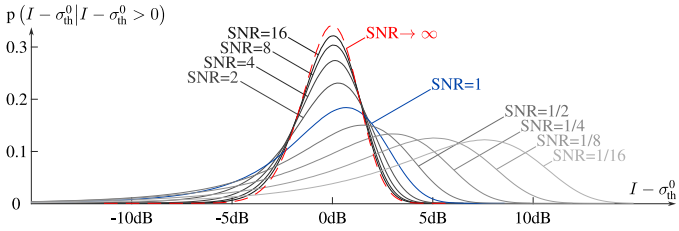


Figure 6. After compensation of the shift due to the thermal noise floor, corrected intensities follow a different statistical distribution. Here, the probability density function of the corrected intensities is represented, in log scale (dB), for different SNR values (ratios of the reflectivity and the thermal noise floor $\boldsymbol{\sigma}^0_{\text{th}}$). Compared to the Fisher-Tippett distribution followed in the absence of thermal noise correction (red dashed curve, SNR $\to \infty$), the shape of the distribution is strongly modified by the correction, preventing from directly applying despeckling methods to corrected images.

second speckled image. In the case of single-polarization images, the SAR2SAR method [27] proposes a strategy to compensate for these changes and consider the content of the second image unchanged, except for the speckle. In this paper, we propose to rely on the SAR2SAR training strategy originally developed for speckle removal of single-channel SLC images and we extend this method to jointly process the polarimetric channels and account for both thermal noise and speckle.

Leveraging temporal information to learn how to despeckle is possible only provided that part of the spatial structures in the images remain stable. This is the case of land areas (topographic structures, patches of woods, fields, roads, human constructions, and infrastructures are generally persistent) but not true of sea ice and the ocean. We address this issue by training our network on images of land and open water areas and then transferring this knowledge to process sea ice images at inference time.

Training procedure: Following the SAR2SAR method [27], the network is trained in two phases, A and B. The aim of phase A is to obtain a despeckling method that can be used to compensate for changes between images used in phase B. It uses images with synthetic speckle. Phase B then fine-tunes the network on actual Sentinel-1 images in order to optimally adapt to the spatial correlations of speckle and the content of these images.

In our experiments, we produced speckle-free ground-truth images by applying a multi-temporal polarimetric filtering method, RABASAR [38], to a stack of 17 RADARSAT-2 images. We then used the 2×2 complex-valued covariance matrix between the HH and HV channels and actual maps of $\boldsymbol{\sigma}^0_{\text{HH}_{\text{th}}}$ and $\boldsymbol{\sigma}^0_{\text{HV}_{\text{th}}}$ from Sentinel-1 EW GRDM to simulate, using the model described in Section II-A and Figure 3, speckled data. To improve the robustness to various SNRs, we used SNR values drawn uniformly in the range $[\frac{1}{10}, 10]$ (i.e., the level of $\boldsymbol{\sigma}^0_{\text{th}}$ was adjusted to cover cases of very low reflectivities with respect to the thermal noise floor as well as more favorable cases). For phase B, we selected images from two land-water areas near the mouth of the river Ob in Russia, in the late summer months of 2017 to 2019 (same orbit number and pass direction). The images have been co-registered and despeckled using the network of phase A in order to compute the radiometric changes occurring between pairs of overlapping images.

Network forward pass: We provide as an input to the network, for each polarization channel, the calibrated images before (diag$(1/\widetilde{\boldsymbol{a}})\widetilde{\boldsymbol{d}}$) and after ($\widetilde{\boldsymbol{d}}_{\text{c}}$) thermal noise subtraction, and the thermal noise map ($\widetilde{\boldsymbol{\sigma}}^0_{\text{th}}$), all in log-scale. The network outputs a two-channels image corresponding to estimates of the reflectivities $\widetilde{\boldsymbol{\sigma}}^0_{\text{HH}}$ and $\widetilde{\boldsymbol{\sigma}}^0_{\text{HV}}$ of the scene, in log-scale. The use of the log scale has two positive effects: it compresses the high-dynamic range of SAR images and stabilizes the variance of speckle.

By combining the corrected image, uncorrected image, and thermal noise map in input, the network is aware of the variance of fluctuations, and the bias term, and can readily use the unbiased image to produce the output image. Since we provide images in log-scale, the thermal noise component can not be recovered by the network by a simple subtraction. We prefer to also provide this information directly as an input.

We connect the input and the output through a U-Net architecture [39] similar to the one used in SAR2SAR [27], with a residual connection between the corrected input and the output, i.e., the network actually only has to estimate the noise that should be removed from the corrected image in order to get the reflectivities $\widetilde{\boldsymbol{\sigma}}^0_{\text{HH}}$ and $\widetilde{\boldsymbol{\sigma}}^0_{\text{HV}}$.

Loss function: The network is trained using pairs of speckled images captured at two different dates. In phase A, we use two independent speckle draws to simulate those images. In phase B, the two images correspond to actual Sentinel-1 EW GRDM images and the second image is modified to compensate for the changes that occurred with respect to the first image: the differences $\log \widetilde{\boldsymbol{\sigma}}^0_{\text{HH}\,2} - \log \widetilde{\boldsymbol{\sigma}}^0_{\text{HH}\,1}$ and $\log \widetilde{\boldsymbol{\sigma}}^0_{\text{HV}\,2} - \log \widetilde{\boldsymbol{\sigma}}^0_{\text{HV}\,1}$, estimated by the network trained in phase A, are subtracted to the second image $\log \widetilde{\boldsymbol{d}}_2$, leading to a change-compensated image $\widetilde{\boldsymbol{d}}^\star_2$.

The same loss function is used for the two phases of the training:

$$\mathcal{L}_{\boldsymbol{\theta}}(\widetilde{\boldsymbol{d}}_1, \widetilde{\boldsymbol{d}}_2) = -\log \mathrm{p}\big(\, \widetilde{\boldsymbol{d}}^\star_2 \,\big|\, f_{\boldsymbol{\theta}}\big[\mathrm{diag}(1/\widetilde{\boldsymbol{a}})\widetilde{\boldsymbol{d}}_1, \widetilde{\boldsymbol{d}}_{\text{c}_1}, \widetilde{\boldsymbol{\sigma}}^0_{\text{th}}\big]\,\big), \tag{12}$$

where $\boldsymbol{\theta}$ denotes the vector of parameters of the deep neural network (i.e., the weights and bias coefficients for all layers) and $f_{\boldsymbol{\theta}}[\cdot]$ represents a forward pass through the network. Minimizing the loss in (12) corresponds to finding the network weights that lead to an output that maximizes the log-likelihood with respect to the change-compensated second image. The likelihood in (12) does not admit a simple expression due to the combination of linear and non-linear operations in the model of equation (4). We consider a simple separable approximation to this likelihood by neglecting spatial and cross-polarization channel correlations, and by approximating the mixture of spatially-correlated intensities by a gamma

Table I
TRAINING HYPERPARAMETERS

	Synthetic speckle (phase A)	Actual speckle (phase B)
# stacks	1	2
# images	17	20
avg images/stack	17	10
patch size	256×256	256×256
batch size	12	12
# patches	5304	11 640
# batches	442	570
# epochs	30	30
learning rate	10^{-3}	10^{-3}
	10^{-4} after 10 epochs	10^{-4} after 10 epochs
	10^{-5} after 20 epochs	10^{-5} after 20 epochs

distribution [40]. This leads to the following expression:

$$-\log \mathrm{p}\big(\widetilde{\boldsymbol{d}} \,\big|\, \widetilde{\boldsymbol{\sigma}}^0 + \widetilde{\boldsymbol{\sigma}}^0_{\mathrm{th}}\big) \approx \sum_n L_n \Bigg[\log\big(\sigma^0_{\mathrm{HH}_n} + \sigma^0_{\mathrm{HH}_{\mathrm{th}\,n}}\big) + \frac{\widetilde{d}_{\mathrm{HH}_n}}{\sigma^0_{\mathrm{HH}_n} + \sigma^0_{\mathrm{HH}_{\mathrm{th}\,n}}} + \log\big(\sigma^0_{\mathrm{HV}_n} + \sigma^0_{\mathrm{HV}_{\mathrm{th}\,n}}\big) + \frac{\widetilde{d}_{\mathrm{HV}_n}}{\sigma^0_{\mathrm{HV}_n} + \sigma^0_{\mathrm{HV}_{\mathrm{th}\,n}}} \Bigg], \tag{13}$$

where L_n corresponds to the equivalent number of looks at pixel n ($L_n \approx 15$ in the first subswath, $L_n \approx 10$ elsewhere), and a constant irrelevant to the minimization problem has been dropped. This approximation, essential to obtain a tractable loss function to train the network, is sufficient to learn a meaningful restoration network. As shown in [27], using a crude approximation of the true likelihood (a Gaussian rather than Fisher-Tippett distribution in the single-polarization and single-look case considered in [27]) still allows to train the network, at the price of a reduced statistical efficiency (i.e., sub-optimal use of the training data).

The details of the training parameters are given in Table I. Our model is available through a Git repository[4].

III. EXPERIMENTS

A. Quantitative evaluation using images with simulated speckle and thermal noise

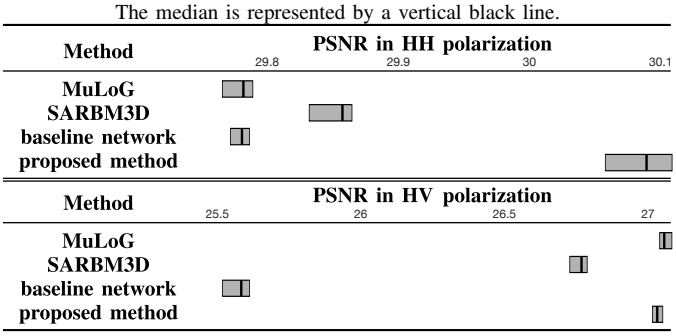
We first perform an evaluation of the proposed approach on simulated Sentinel-1 EW GRDM data. The purpose of these simulations is to conduct an evaluation with known ground truth to better analyze the artifacts introduced by the different methods and perform a quantitative comparison.

Computation of a ground-truth image: We used a C-band fine quad-polarization image from the Radarsat-2 satellite, which offers higher-resolution images with a lower thermal noise level $\boldsymbol{\sigma}^0_{\mathrm{th}}$ than Sentinel-1. Speckle noise was suppressed in two steps: first, by multilooking to reduce the spatial resolution to that of a Sentinel-1 EW GRDM image (leading to images with an equivalent number of looks of about 30), and second by applying a multi-channel despeckling filter (MuLoG [41]), to remove the residual fluctuations. The reflectivities $\boldsymbol{\sigma}^0_{\mathrm{HH}}$ and $\boldsymbol{\sigma}^0_{\mathrm{HV}}$ corresponding to the first two elements of the diagonal of the full-polarimetric covariance matrices in the resulting image form our ground-truth, see left-hand side of section 1 of Figure 7.

[4]https://gitlab.telecom-paris.fr/ring/joint-despeckling-and-thermal-noise-compensation.git

Table II
QUANTITATIVE EVALUATION ON SYNTHETIC IMAGES CORRUPTED BY SPECKLE AND THERMAL NOISE

The grey rectangles extend from the first to the third quartile of the PSNR values (PSNR is computed for 20 noise realizations). The median is represented by a vertical black line.



Simulation of noisy images: We followed the generative model described in Figure 3. The thermal noise floor $\boldsymbol{\sigma}^0_{\mathrm{th}}$ has been set by extracting a patch of same size as the ground truth image from an actual Sentinel-1 EW GRDM NESZ map, at a region overlapping the first two sub-swaths (in order to analyze the impact of this discontinuity). The average SNR $\langle \widetilde{\sigma}^0_{\mathrm{HH}} \rangle / \langle \widetilde{\sigma}^0_{\mathrm{HH}_{\mathrm{th}}} \rangle$ has been set to 10, to fit typical values observed in Sentinel-1 images. Uncorrelated speckle has been synthesized, with an equivalent number of looks $L = 15$ and $L = 10$ for the first and second sub-swath, respectively, see right-hand side of section 1 of Figure 7. Note that, to simplify the simulation, we did not model the spatial correlations of noise due to the linear operators $\mathbf{H}$ and $\mathbf{S}$. Therefore, we applied our networks directly after Phase A of the training (i.e., training on images corrupted by synthetic speckle). Twenty different noise realizations were produced for the same ground truth image.

Quantitative analysis: Table II compares the Peak Signal to Noise Ratio (PSNR) computed between estimated and ground truth values of $\log \widetilde{\boldsymbol{\sigma}}^0$, in HH and HV polarizations, for 4 restoration strategies:

1) despeckling by application of MuLoG [41] with the denoising algorithm BM3D [42], followed by the subtraction of the thermal noise floor $\boldsymbol{\sigma}^0_{\mathrm{th}}$;
2) despeckling by application of SARBM3D [43], followed by the subtraction of the thermal noise floor $\boldsymbol{\sigma}^0_{\mathrm{th}}$;
3) despeckling by a baseline network described in [37], trained to despeckle jointly the HH and HV channels using a similar U-Net architecture, followed by the subtraction of the thermal noise floor $\boldsymbol{\sigma}^0_{\mathrm{th}}$;
4) the proposed approach: subtraction of the thermal noise floor $\boldsymbol{\sigma}^0_{\mathrm{th}}$ followed by our NESZ-aware restoration network introduced in Section II-C.

In Table II, we represent by a gray rectangle the half of PSNR values included between the first and third quartiles. The median value is indicated by a vertical bar. The proposed method outperforms other methods in the HH channel and

performs comparably to MuLoG in the HV channel. Note that the gap between the performance of standard methods (MuLoG and SARBM3D) and self-supervised networks (baseline and proposed) increases when processing real data due to the speckle correlations which are not accounted for by standard methods but are addressed by the fine-tuning step during phase B of our training.

Section 2 of Figure 7 shows the images obtained by the different methods on one of the noise realizations. Zooms in 3 areas are provided. The dual-polarization information is represented in color using the following convention: green and blue channels contain $\log \widetilde{\boldsymbol{\sigma}}_{\mathrm{HH}}$ (the reflectivity of the HH polarization in log-scale), and the red channel contains $\log \widetilde{\boldsymbol{\sigma}}_{\mathrm{HV}}$ (the reflectivity of the HV polarization in log-scale). Each channel is normalized to cover the range from the 5th to the 95th percentile. Close inspection of the images, in particular in the 3 zoomed-in areas, shows that fine details are better recovered with the proposed approach.

B. Qualitative evaluation on an S1 EW GRDM image

We first validate on a Sentinel-1 EW GRDM image our theoretical discussion of the impact on the noise statistics of removing the thermal noise floor. Figure 8 shows the histograms before and after removing the offset due to thermal noise. Empirical histograms were obtained by selecting all pixels in the first sub-swath of the HV channel with an estimated SNR equal to 0dB$\pm$0.41dB (corresponding to $\pm 10\%$ in linear scale, which represents for our image approximately 12% of the pixels of the first sub-swath). In order to compensate for variations of the reflectivity and thermal noise levels among selected pixels, observed noisy values are normalized: we compute the empirical histogram $\mathrm{p}\left(\left[\frac{\mathrm{diag}(1/\widetilde{\boldsymbol{a}}_{\mathrm{HV}})\widetilde{\boldsymbol{d}}_{\mathrm{HV}}}{\widetilde{\boldsymbol{\sigma}}^0_{\mathrm{HV}}+\widetilde{\boldsymbol{\sigma}}^0_{\mathrm{HV_{th}}}}\right]_n\right)$ corresponding to calibrated values *before* thermal floor removal, in blue, and histogram $\mathrm{p}\left(\left[\frac{\mathrm{diag}(1/\widetilde{\boldsymbol{a}}_{\mathrm{HV}})\widetilde{\boldsymbol{d}}_{\mathrm{HV}}-\widetilde{\boldsymbol{\sigma}}^0_{\mathrm{HV_{th}}}}{\widetilde{\boldsymbol{\sigma}}^0_{\mathrm{HV}}}\right]_n\right)$ corresponding to calibrated values *after* thermal floor removal, in yellow, for all the pixels n that satisfy the SNR constraint: $-0.41\mathrm{dB} \leq \mathrm{SNR}_{\mathrm{HV}n} = \frac{\widetilde{\sigma}^0_{\mathrm{HV}n}}{\widetilde{\sigma}^0_{\mathrm{HV_{th}}n}} \leq 0.41\mathrm{dB}$, where MuLoG's despeckling result [41] provides an estimate of the values $\widetilde{\boldsymbol{\sigma}}^0_{\mathrm{HV}}$. Superimposed to these empirical histograms drawn as bar graphs, the theoretical curves are represented with dashed lines. Despite the estimation of $\widetilde{\boldsymbol{\sigma}}^0_{\mathrm{HV}}$ being imperfect in the pixel selection and intensity normalization steps, a good match between the theoretical and experimental histograms is observed. As expected from our theoretical analysis, removing the thermal floor has a very significant impact on the noise distribution (blue and yellow distributions are quite different), justifying the development of a dedicated restoration method.

We compare in Figure 9 the restoration of an actual Sentinel-1 EW GRDM image. We show only the HV channel which is the one with the worst signal-to-noise ratio. We selected an area overlapping the first and second sub-swaths to illustrate the impact of the discontinuity in the thermal noise floor $\boldsymbol{\sigma}^0_{\mathrm{th}}$ (identified by a red vertical dashed line in Figure 9(a)). After subtraction of the thermal noise offset, this discontinuity is much less visible (what remains is the difference in the variance of fluctuations in each sub-swath), see Figure 9(b). Yet, all standard despeckling techniques are applicable only to the original image Figure 9(a). The results shown in Figure 9(c) to Figure 9(e) for MuLoG [41], SARBM3D [43], and our deep neural network baseline [37] all display artifacts at the boundary between sub-swaths, as pointed out by the red arrows. In contrast, the result obtained with the proposed method shown in Figure 9(f) does not suffer from this problem.

C. Impact of the thermal noise suppression and despeckling on downstream applications: illustration on ice classification

There is a large number of downstream applications of SAR remote sensing that can benefit from improved speckle reduction in Sentinel-1 wide-swath imagery. As an example case, we here choose operational sea ice monitoring, which is essential to support maritime navigation in the polar regions and to ensure the safety of Arctic offshore operations [44]. Sea ice conditions are routinely mapped by national ice services around the world and the resulting information is distributed in the form of ice charts. While operational ice chart production is at present still performed manually, multiple studies have investigated the (semi-)automated separation of sea ice and open water as well as the classification of different sea ice types, using both deep-learning approaches and statistical methods [45].

Here, we choose a pixel-wise classification algorithm introduced by Lohse et al. [13], which is currently being tested and evaluated in an operational setting at the Norwegian Ice Service. The method uses the local incidence angle together with both HH and HV backscatter intensities of the Sentinel-1 EW GRDM product. It accounts for the well-known effect of class-dependent backscatter variation with the incidence angle by assuming a linearly variable mean vector for each individual class distribution. In this study, we use a version of the classifier that was specifically trained for the area around Belgica Bank in Western Fram Strait to support navigation during a research cruise into that area in April and May 2022. The algorithm distinguishes four ice classes (*Open Water/New Ice*, *Young Ice*, *Level Ice*, *Deformed Ice*) which are relevant for tactical navigation in areas of high sea ice concentration. Details about the class definitions and training data selection can be found in [46], [47].

For the demonstration of how the proposed speckle reduction affects the classification results, we selected four images from two Sentinel-1 overpasses on May 2nd and May 3rd 2022. We separately applied seven different processing chains for speckle reduction and thermal noise removal: (i) no speckle reduction, multilooking with two different window sizes (ii) 9×9 and (iii) 21×21, the three state-of-the-art despeckling methods: (iv) Mulog with BM3D [41], (v) SAR-BM3D [43], (vi) the dual-polarization deep neural network [37], and finally (vii) the method proposed in this paper. We classify the output image from each method and geocode the results to a Polar Stereographic projection (epsg: 3996) with 40 m pixel spacing.

Figure 10 shows an overview of the area of interest, with a false-color intensity image from one Sentinel-1 overpass

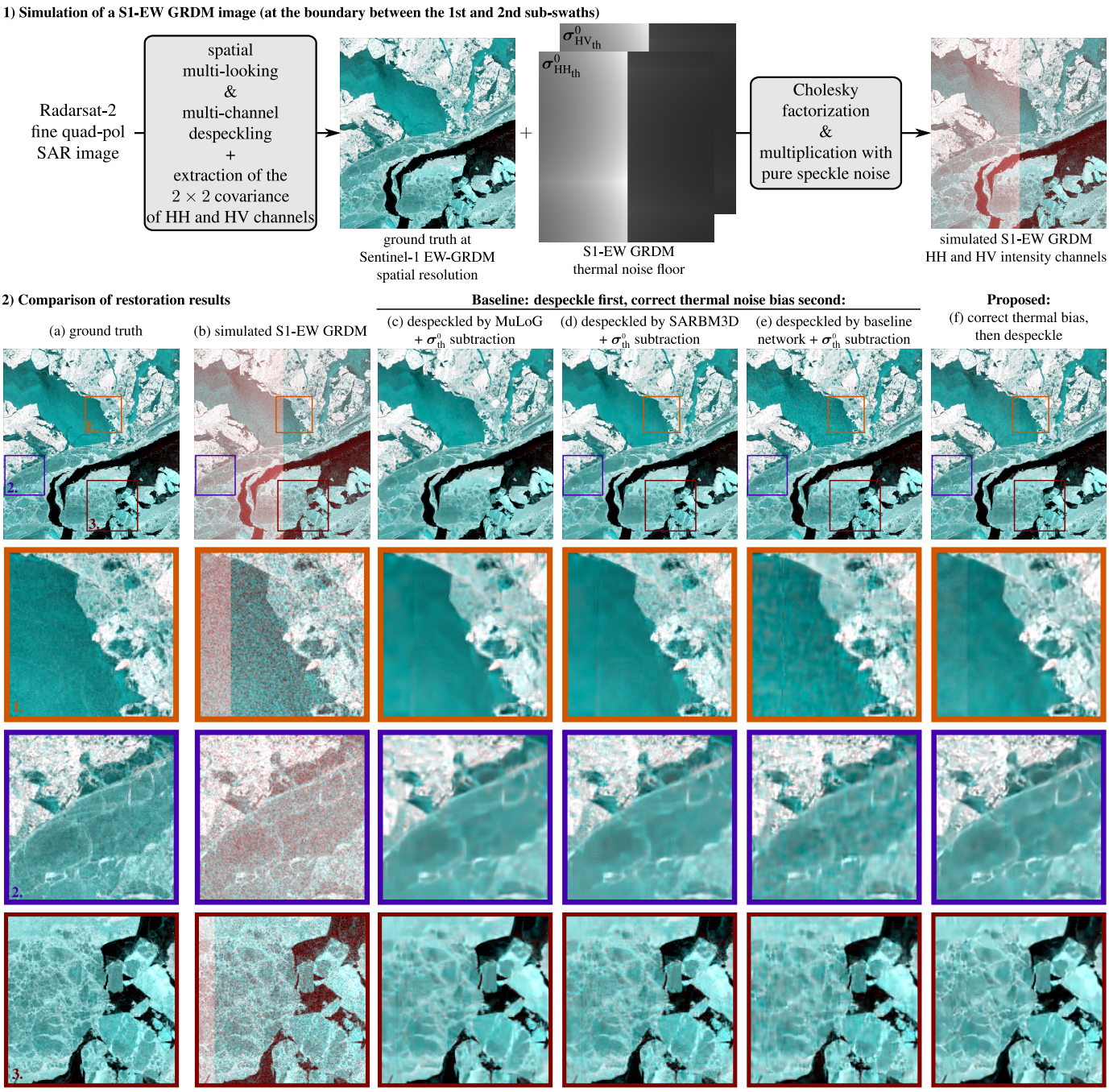


Figure 7. Evaluation of restoration methods on a simulation of Sentinel-1 EW GRDM data: (a) ground-truth image obtained from a Radarsat-2 image (©Government of Canada); (b) simulated image using an actual Sentinel-1 thermal noise map σ^0_{th} overlapping the first and second sub-swaths; (c-e) methods that despeckle first, then subtract σ^0_{th}; (f) proposed approach.

on the left and the corresponding classification result on the right. Regions of interest (ROIs) for the different ice types that are used for quantitative evaluation of the classification results from the different processing chains are overlaid on the intensity image. However, such a quantitative assessment is not straightforward, since the main improvement is expected in regions either highly affected by thermal noise or corresponding to detailed spatial variations in the ice pack. Hence, "standard" ROIs that are often drawn in extended homogeneous areas or derived from coarse polygons of operational ice charts are not suitable to demonstrate the improvements. We have therefore carefully selected small validation ROIs (5×5 and 9×9 pixels, corresponding to 200×200 and 360×360 meters, respectively) along or in the close vicinity of the sub-swath boundary between EW1 and EW2. This boundary is often challenging for the correction of thermal noise, and the small size of the ROIs enables us to capture small-scale variations in the classification results. While this approach is suitable to demonstrate the improvements from our processing method, it should be kept in mind that the absolute numbers reported for

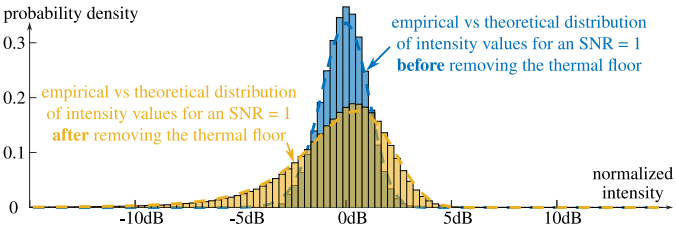


Figure 8. Empirical verification of the impact of thermal floor removal on the statistics of noise: two empirical histograms computed on the first sub-swath of a Sentinel-1 EW GRDM image are compared with the theoretical distributions drawn in dashed lines for pixels with an estimated SNR of 0dB±0.41dB, before (in blue) and after (in yellow) removal of the thermal noise floor.

Table III
CLASSIFICATION ACCURACY (IN %)

	OW / New Ice	Young Ice	Level Ice	Deformed Ice	Average per-class
ML 1x1	71.8	43.2	80.8	82.3	69.5
ML 9x9	88.2	76.1	83.8	98.1	86.6
ML 21x21	44.3	60.4	70.0	98.1	68.2
MuLoG	94.7	71.7	75.3	95.5	84.3
SARBM3D	93.6	69.7	73.0	95.0	82.8
Baseline	84.8	62.3	61.5	94.1	75.7
Proposed	96.4	78.3	86.8	97.4	89.7

classification accuracy (CA) are pessimistic since we focused on small validation areas where classification is the most challenging (to emphasize the differences between the tested methods).

Table III summarizes the results for the per-class and average per-class CA from all tested approaches, with the highest scores highlighted in green and particularly low scores highlighted in red. Our proposed method achieves the highest score for three of the four classes as well as the best average per-class CA. Notably, the two multilooking approaches ML 9×9 and ML 21×21 achieve the highest scores for the *Deformed Ice*. We attribute this to the rather large spatial areas of *Deformed Ice*, which makes them less susceptible to errors introduced by the spatial averaging in the multi-looking approaches. However, the ice types in rather small spatial areas, such as *Open Water/New Ice* and *Young Ice* within lead systems, score considerably lower for ML 9×9 and ML 21×21. Especially the large multilooking window performs poorly with regard to the spatial detail of these classes. The state-of-the-art despeckling methods (MuLoG and SARBM3D) achieve overall very similar results to each other and score slightly lower than our proposed method (5 and 7% for average per-class CA). This difference is most likely caused by small variations in maintaining spatial detail as well as their reduced performance along the sub-swath boundary. We discuss this further in the qualitative assessment of the results following paragraph.

Zoomed-in close-up comparisons for qualitative evaluation of the results are presented in Figure 11 and Figure 12. While Figure 11 shows six out of the seven tested approaches, Figure 12 only includes the three highest scoring methods in terms of average per-class CA.

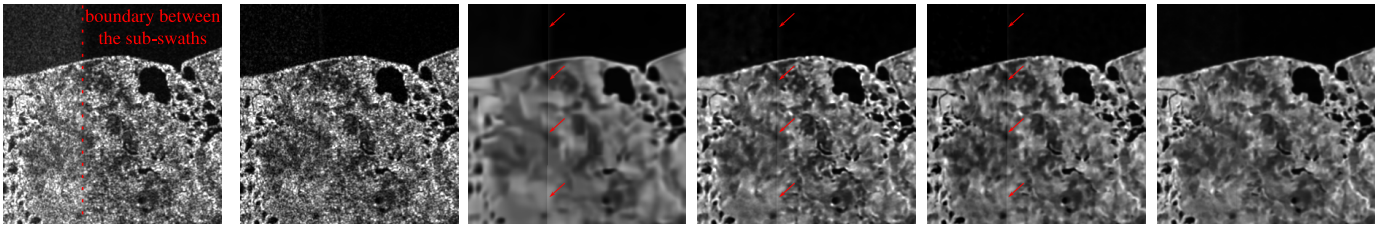
The qualitative comparison is in good agreement with the quantitative CA results and clearly shows that the baseline methods as well as our new proposed approach outperform the multilooking approaches with respect to spatial detail. The classification results appear significantly smoother than the ML 1×1 results, while at the same time maintaining small-scale structures such as narrow leads much better than both ML 9×9 and ML 21×21. Accurate mapping of such structures is critical for tactical navigation and routing support for ice-going vessels. Large homogeneous areas on the other hand, which are mostly found for the *Deformed Ice* class, are overall well delineated by the multilooking as well as the deep-learning approaches. Furthermore, Figure 11 shows that the classification results from the baseline methods all suffer from artifacts at the sub-swath boundaries. These artifacts are identified by red arrows in the figure. Note that after the polar stereographic projection, the boundary between sub-swaths is no longer vertical. Our proposed method is able to remove these artifacts, leading to an overall smoother and more consistent classification result where the inter-sub-swath boundary is not visible, which is preferable for further use in either tactical navigation or data assimilation in numerical models for sea ice forecasts.

IV. CONCLUSION

The analysis of Sentinel-1 images of the Arctic is difficult due to speckle fluctuations and because of the impact of thermal noise, in particular in the cross-polarization channel. Previous methods in the literature addressed only one of these two aspects. Combining thermal noise correction and despeckling can be challenging, especially at the boundary between sub-swaths. We have developed a statistical model describing the impact of both speckle and thermal noise in polarimetric images. Using this model, we have shown that standard despeckling techniques must be applied to images contaminated by the thermal noise bias otherwise the noise distribution strongly departs from Goodman's speckle model. Yet, correcting this bias *a posteriori* leads to artifacts in regions where the thermal noise floor varies strongly, such as between sub-swaths of Sentinel-1 EW images. To prevent these artifacts, it is necessary to process images corrected for the thermal noise bias. A specific method must then be developed to account for the modified statistics of the speckle and thermal noise fluctuations. We proposed a self-supervised approach based on a deep-neural network informed by the map of the thermal noise floor. This strategy is shown to better restore the images and reduce artifacts at the boundary between sub-swaths. A better estimation of polarimetric reflectivities is crucial in downstream applications such as ice classification. We found a strong improvement with respect to multilooking and avoided the accumulation of classification errors at the boundary between sub-swaths with our new approach.

ACKNOWLEDGEMENTS

This work has been supported by the French-Norwegian program NFR-PHC Aurora under project COSMIC (Advanced processing of SAR images for the Arctic), RCN grant nr. 331935, Campus France nr. 48345RK, and the Visual Intelligence Centre for Research-based Innovation funded by the



(a) Sentinel-1 EW GRDM image (HV polarization) (b) image after subtracting $\boldsymbol{\sigma}^0_{\text{th}}$ (c) despeckled by MuLoG + $\boldsymbol{\sigma}^0_{\text{th}}$ subtraction (d) despeckled by SARBM3D + $\boldsymbol{\sigma}^0_{\text{th}}$ subtraction (e) despeckled by baseline network + $\boldsymbol{\sigma}^0_{\text{th}}$ subtraction (f) **proposed method**

Figure 9. Application to an actual Sentinel-1 EW GRDM image (HV polarization channel) of an area in Northern Russia (©ESA): (a) original image; (b) after correcting for the thermal noise bias; (c-e) restored image by despeckling image (a) and subtracting the thermal *a posteriori*; (f) proposed method: restoration of image (b) with a network aware of the thermal noise map $\boldsymbol{\sigma}^0_{\text{th}}$. Artifacts pointed by red arrows are visible at the boundary between the sub-swaths indicated by the vertical dashed line in (a) for methods (c-e) and are almost absent of the result (f).

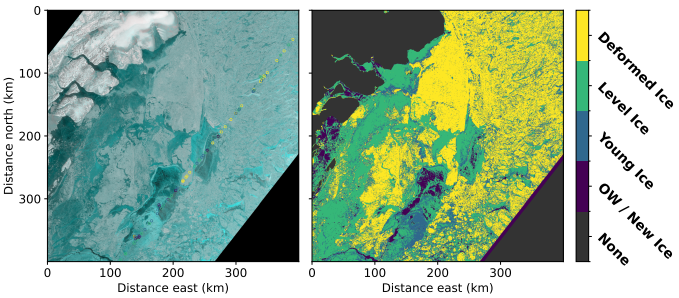


Figure 10. False-color intensity image (left, RGB: HV, HH, HH) and corresponding classification result (right) for the area of interest after applying our proposed speckle reduction method. Validation ROIs for different ice types are overlaid on the intensity image.

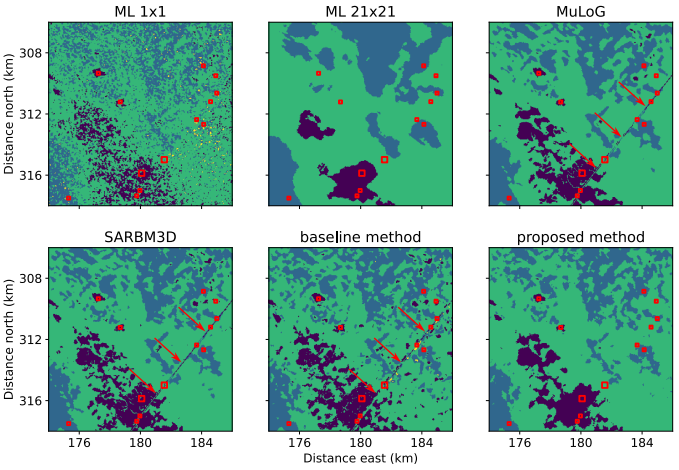


Figure 11. Close-up comparison of ice type classification results (same colorbar as in Figure 10) after different despeckling methods. Red rectangles outline validation ROIs used for quantitative evaluation. The results show that our proposed method significantly reduces speckle effects in the classification result while maintaining spatial details. Furthermore, swath boundary artifacts pointed by the red arrows in the results from MuLoG, SARBM3D, and the baseline deep neural network, are successfully suppressed.

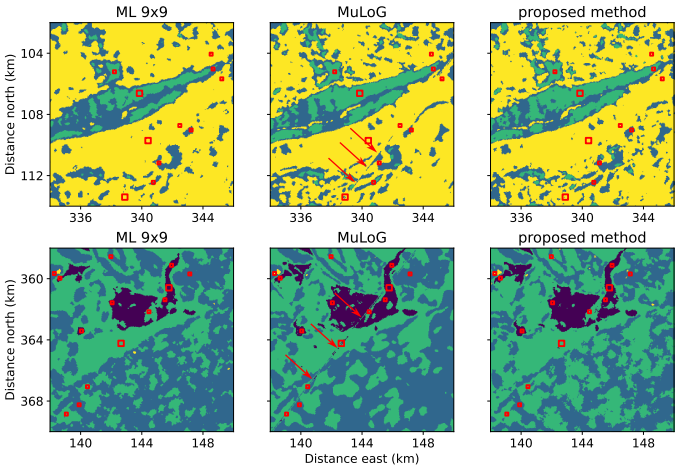


Figure 12. Two examples (top and bottom row) showing close-up comparisons of ice type classification results (same colorbar as in Figure 10) after despeckling with the three highest scoring methods according to Table III. Red rectangles outline validation ROIs used for quantitative evaluation. Differences are less pronounced than in Figure 11, yet both deep-learning approaches maintain finer spatial detail than the traditional ML and swath boundary artifacts (pointed out by the red arrows) are still visible in the results from MuLoG.

Research Council of Norway (RCN grant nr. 309439), and the NATALIE project funded by the European Union Horizon Europe Climate research and innovation program under grant agreement nr. 101112859.

F. Tupin, E. Dalsasso, and L. Denis would also like to thank the French National Research Agency (ANR) and the Agence de l'Innovation de Défence (AID) for funding under ASTRAL project ANR-21-ASTR-0011.

The authors are grateful to the Editors and Reviewers for their comments that helped improve the manuscript.

APPENDIX A
EXTENSION OF THE DENOISING APPROACH TO FULLY-POLARIMETRIC IMAGES

The description of the method has been given in the case of dual polarimetric images, which covers the important practical case of Sentinel-1 data. The extension to images from sensors with full-polarimetry capabilities is discussed here.

When taking into account the thermal noise, the polarimetric covariance matrix at pixel n of a fully polarimetric monostatic system is given by equation (14) displayed at the top of page 12. The Cholesky factorization of this matrix $\boldsymbol{\Sigma}_{\text{tot}_n} = \mathbf{M}_{\text{tot}_n}\mathbf{M}^*_{\text{tot}_n}$, useful to generate synthetic images for the first step of the training of the restoration network, is

$$\mathbf{\Sigma}_{\text{tot}_n} = \begin{pmatrix} \sigma^0_{\text{HH}_n} + \sigma^0_{\text{HH}_{\text{th}\,n}} & \sqrt{\sigma^0_{\text{HH}_n}\sigma^0_{\text{HV}_n}}\rho_{\text{HH HV}_n}e^{j\beta_{\text{HH HV}_n}} & \sqrt{\sigma^0_{\text{HH}_n}\sigma^0_{\text{VV}_n}}\rho_{\text{HH VV}_n}e^{j\beta_{\text{HH VV}_n}} \\ \sqrt{\sigma^0_{\text{HH}_n}\sigma^0_{\text{HV}_n}}\rho_{\text{HH HV}_n}e^{-j\beta_{\text{HH HV}_n}} & \sigma^0_{\text{HV}_n} + \sigma^0_{\text{HV}_{\text{th}\,n}} & \sqrt{\sigma^0_{\text{HV}_n}\sigma^0_{\text{VV}_n}}\rho_{\text{HV VV}_n}e^{j\beta_{\text{HV VV}_n}} \\ \sqrt{\sigma^0_{\text{HH}_n}\sigma^0_{\text{VV}_n}}\rho_{\text{HH VV}_n}e^{-j\beta_{\text{HH VV}_n}} & \sqrt{\sigma^0_{\text{HV}_n}\sigma^0_{\text{VV}_n}}\rho_{\text{HV VV}_n}e^{-j\beta_{\text{HV VV}_n}} & \sigma^0_{\text{VV}_n} + \sigma^0_{\text{VV}_{\text{th}\,n}} \end{pmatrix} \tag{14}$$

defined by:

$$[\mathbf{M}_{\text{tot}_n}]_{j,j} = \sqrt{[\mathbf{\Sigma}_{\text{tot}_n}]_{j,j} - \sum_{k=1}^{j-1}\left|[\mathbf{M}_{\text{tot}_n}]_{j,k}\right|^2}, \tag{15}$$

and

$$[\mathbf{M}_{\text{tot}_n}]_{i,j} = \frac{1}{[\mathbf{M}_{\text{tot}_n}]_{j,j}}\left([\mathbf{\Sigma}_{\text{tot}_n}]_{i,j} - \sum_{k=1}^{j-1}[\mathbf{M}_{\text{tot}_n}]^*_{j,k}[\mathbf{M}_{\text{tot}_n}]_{i,k}\right) \tag{16}$$

for $i > j$. Equations (15) and (16) boil down to equation (2) in the dual-polarization case, and slightly more complex expressions in the full-polarimetric case (the expansion is not reproduced here due to space constraints). Given that the Cholesky decomposition has to be computed for each pixel of the image, the expressions (15) and (16) are much more efficient in practice for array-oriented languages such as Python or Matlab when applied to the whole image at once, compared to repetitively calling an external function for Cholesky factorization at each pixel.

Beyond the modification of the model and the simulation of corrupted images, the network inputs and outputs must be adjusted. We described in Section II-C that our network takes, for each polarization channel, the calibrated images before and after thermal noise subtraction, as well as the thermal noise map. The input dimension thus corresponds to 6 channels for dual-polarization images and 9 channels for fully-polarimetric images. In terms of output, in dual-polarization a two-channels image corresponding to estimates of the reflectivities $\widetilde{\boldsymbol{\sigma}}^0_{\text{HH}}$ and $\widetilde{\boldsymbol{\sigma}}^0_{\text{HV}}$ of the scene is produced, while in the case of fully-polarimetric imaging a three-channels image would be estimated ($\widetilde{\boldsymbol{\sigma}}^0_{\text{HH}}$, $\widetilde{\boldsymbol{\sigma}}^0_{\text{HV}}$ and $\widetilde{\boldsymbol{\sigma}}^0_{\text{VV}}$).

REFERENCES

[1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 1, pp. 6–43, 2013.

[2] S. Cloude, *Polarisation: applications in remote sensing*. OUP Oxford, 2009.

[3] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse problems*, vol. 14, no. 4, p. R1, 1998.

[4] I. Hajnsek and Y.-L. Desnos, *Polarimetric Synthetic Aperture Radar: Principles and Application*. Springer Nature, 2021, vol. 25.

[5] M. Rantanen, A. Y. Karpechko, A. Lipponen, K. Nordling, O. Hyvärinen, K. Ruosteenoja, T. Vihma, and A. Laaksonen, "The Arctic has warmed nearly four times faster than the globe since 1979," *Communications Earth & Environment*, vol. 3, no. 1, p. 168, 2022.

[6] N. Zakhvatkina, V. Smirnov, and I. Bychkova, "Satellite SAR data-based sea ice classification: An overview," *Geosciences*, vol. 9, no. 4, p. 152, 2019.

[7] W. Dierking, "Mapping of Different Sea Ice Regimes Using Images From Sentinel-1 and ALOS Synthetic Aperture Radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1045–1058, 2010.

[8] E. Khachatrian, W. Dierking, S. Chlaily, T. Eltoft, F. Dinessen, N. Hughes, and A. Marinoni, "Sar and passive microwave fusion scheme: A test case on sentinel-1/amsr-2 for sea ice classification," *Geophysical Research Letters*, vol. 50, no. 4, p. e2022GL102083, 2023.

[9] F. De Zan and A. M. Guarnieri, "TOPSAR: Terrain observation by progressive scans," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2352–2360, 2006.

[10] "L1 GRD Products," https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-1-algorithms/ground-range-detected, accessed: 2022-06-14.

[11] J.-W. Park, A. A. Korosov, M. Babiker, J.-S. Won, M. W. Hansen, and H.-C. Kim, "Classification of sea ice types in Sentinel-1 synthetic aperture radar images," *The Cryosphere*, vol. 14, no. 8, pp. 2629–2645, 2020.

[12] M. Mäkynen and J. Karvonen, "Incidence Angle Dependence of First-Year Sea Ice Backscattering Coefficient in Sentinel-1 SAR Imagery Over the Kara Sea," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 11, pp. 6170–6181, 2017.

[13] J. Lohse, A. P. Doulgeris, and W. Dierking, "Mapping sea-ice types from Sentinel-1 considering the surface-type dependent effect of incidence angle," *Annals of Glaciology*, vol. 61, no. 83, pp. 260–270, 2020.

[14] R. Piantanida, N. Miranda, and N. Franceschi, "Thermal denoising of products generated by the S-1 IPF," Jan 2017. [Online]. Available: https://sentinel.esa.int/documents/247904/2142675/Thermal-Denoising-of-Products-Generated-by-Sentinel-1-IPF

[15] J. Karvonen, "Baltic sea ice concentration estimation using SENTINEL-1 SAR and AMSR2 microwave radiometer data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2871–2883, 2017.

[16] J.-W. Park, A. A. Korosov, M. Babiker, S. Sandven, and J.-S. Won, "Efficient thermal noise removal for Sentinel-1 TOPSAR cross-polarization channel," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1555–1565, 2017.

[17] J.-W. Park, J.-S. Won, A. A. Korosov, M. Babiker, and N. Miranda, "Textural noise correction for Sentinel-1 TOPSAR cross-polarization channel images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 4040–4049, 2019.

[18] P. Q. Lee, L. Xu, and D. A. Clausi, "Sentinel-1 additive noise removal from cross-polarization extra-wide TOPSAR with dynamic least-squares," *Remote Sensing of Environment*, vol. 248, p. 111982, 2020.

[19] Y. Sun and X.-M. Li, "Denoising Sentinel-1 Extra-Wide Mode Cross-Polarization Images Over Sea Ice," *IEEE Trans. Geosci. Remote. Sens.*, vol. 59, no. 3, pp. 2116–2131, 2021.

[20] A. Korosov, D. Demchev, N. Miranda, N. Franceschi, and J.-W. Park, "Thermal denoising of cross-polarized Sentinel-1 data in interferometric and extra wide swath modes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2021.

[21] F. Argenti, A. Lapini, T. Bianchi, and L. Alparone, "A tutorial on speckle reduction in synthetic aperture radar images," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 3, pp. 6–35, 2013.

[22] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: The nonlocal paradigm," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 69–78, 2014.

[23] G. Fracastoro, E. Magli, G. Poggi, G. Scarpa, D. Valsesia, and L. Verdoliva, "Deep learning methods for synthetic aperture radar image despeckling: An overview of trends and perspectives," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 2, pp. 29–51, 2021.

[24] B. Rasti, Y. Chang, E. Dalsasso, L. Denis, and P. Ghamisi, "Image restoration for remote sensing: Overview and toolbox," *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, no. 2, pp. 201–230, 2021.

[25] Y. Wang, C. M. Albrecht, N. A. A. Braham, L. Mou, and X. X. Zhu, "Self-supervised learning in remote sensing: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, no. 4, pp. 213–247, 2022.

[26] E. Dalsasso, L. Denis, M. Muzeau, and F. Tupin, "Self-supervised training strategies for SAR image despeckling with deep neural networks," in *EUSAR 2022; 14th European Conference on Synthetic Aperture Radar*. VDE, 2022, pp. 1–6.

[27] E. Dalsasso, L. Denis, and F. Tupin, "SAR2SAR: A semi-supervised despeckling algorithm for SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4321–4329, 2021.

[28] C. U. Mendes, E. Dalsasso, Y. Zhang, L. Denis, and F. Tupin, "PolSAR2PolSAR: A semi-supervised despeckling algorithm for polarimetric SAR images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 220, pp. 783–798, 2025.

[29] A. B. Molini, D. Valsesia, G. Fracastoro, and E. Magli, "Speckle2Void: Deep self-supervised SAR despeckling with blind-spot convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2021.

[30] E. Dalsasso, L. Denis, and F. Tupin, "As if by magic: self-supervised training of deep despeckling networks with MERLIN," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2021.

[31] J.-S. Lee, K. P. Papathanassiou, T. L. Ainsworth, M. R. Grunes, and A. Reigber, "A new technique for noise filtering of sar interferometric phase images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1456–1465, 1998.

[32] C. López-Martínez and X. Fabregas, "Polarimetric SAR speckle noise model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 10, pp. 2232–2242, 2003.

[33] C. Lopez-Martinez and E. Pottier, "On the extension of multidimensional speckle noise model from single-look to multilook sar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 2, pp. 305–320, 2007.

[34] C. Lopez-Martinez and X. Fabregas, "Model-based polarimetric sar speckle filter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3894–3907, 2008.

[35] I. Meraoumia, E. Dalsasso, L. Denis, R. Abergel, and F. Tupin, "Multitemporal speckle reduction with self-supervised deep neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.

[36] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.

[37] I. Meraoumia, D. Ratha, E. Dalsasso, L. Denis, F. Tupin, and A. Marinoni, "Despeckling of dual-pol GRD Sentinel-1 images in extra-wide mode by deep learning," in *International Geoscience and Remote Sensing Symposium (IGARSS) 2023*, 2023.

[38] W. Zhao, C.-A. Deledalle, L. Denis, H. Maître, J.-M. Nicolas, and F. Tupin, "Ratio-Based Multitemporal SAR Images Denoising: RABASAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 3552–3565, 2019.

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[40] J. W. Goodman, *Speckle phenomena in optics: theory and applications*. Roberts and Company Publishers, 2007.

[41] C. Deledalle, L. Denis, S. Tabti, and F. Tupin, "MuLoG, or How to apply gaussian denoisers to multi-channel SAR speckle reduction ?" *IEEE Transactions on Image Processing*, 2017.

[42] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[43] S. Parrilli, M. Poderico, C. V. Angelino, and L. Verdoliva, "A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 606–616, 2012.

[44] N. Zakhvatkina, V. Smirnov, and I. Bychkova, "Satellite SAR data-based sea ice classification: An overview," *Geosciences*, vol. 9, no. 4, p. 152, 2019.

[45] J. Lohse, "On Automated Classification of Sea Ice Types in SAR Imagery," PhD thesis, UiT The Arctic University of Norway, Tromsø, Norway, Mar 2021, available at https://hdl.handle.net/10037/20606.

[46] J. Lohse, C. Taelman, A. Everett, and N. Hughes, "Fully-automated navigation support for vessels in the Arctic: An application and validation example of ice type mapping during the CIRFA cruise 2022," Mar 2023. [Online]. Available: https://doi.org/10.5281/zenodo.7774480

[47] ——, "Automated Sentinel-1 ice type mapping and in-situ validation during the CIRFA-22 cruise," *Annals of Glaciology*, vol. 65, no. e39, pp. 1–12, 2024.