


Department of Information Engineering and Computer Science

Master's Degree in

Information and Communication Engineering

FINAL DISSERTATION

# ESTIMATION OF SAR QUANTIZATION ERROR: POTENTIALS OF DEEP LEARNING AND CLASSICAL METHODS

Supervisor

Prof. Lorenzo Bruzzone

Student

Daniel Carcereri

Co-Supervisors

Dr. Michele Martone (DLR)

Dr. Francescopaolo Sica (DLR)

Academic year 2019/2020

# Abstract

For spaceborne SAR missions, increasing memory requirements represent an on-going trend, as the deployment of systems using larger bandwidths, covering wider swaths, using multiple polarizations, and employing shorter pulse repetition intervals, strain the available on-board memory and downlink capacity. In this context, smart resource allocation and high-performance quantization techniques represent key research areas.

This master thesis explores the potentials of two novel ideas related to the topic of SAR data volume reduction. Starting from an overview of state-of-the-art quantization techniques, the magnitude of the quantization noise is studied, considering the signal-dependent nature of the phenomena, as well as its propagation from the raw-data domain to that of the focused SAR image. From here, a novel smart resource management concept is investigated. It consists in adaptively allocating quantization resources based on the predicted local error magnitude. To this end, a set of local image statistics has been computed, and then tested for linear correlation with SAR and InSAR performance metrics. Among these, the standard deviation map showed the highest linear correlation, with a Pearson Correlation Coefficient of 0.59.

Furthermore, a second data volume reduction technique is conceptualized, consisting of a deep residual learning-based denoiser, used on ground to compensate the distortions introduced by the on-board, lossy quantization algorithms. Depending on the considered test-site, the implemented architecture achieves average performance gains ranging from 0.4 to 2.1 dB, demonstrating the potentials of the technique.

# List of Figures


1 On the red circle lie all the possible target locations associated to the echo received at time $\tau_0$. The green pseudo-hyperbola identifies the possible target locations given the Doppler shift observed by the platform moving at speed $v_s$. The intersection identifies the remaining forward/backward and left/right ambiguities. . . . . 4
2 Typical SAR acquisition geometry. . . . . 5
3 SAR focusing process. . . . . 7
4 SAR system at instants $t_1$ and $t_2$. The resulting synthetic aperture is equal to halt the antenna aperture in azimuth $L_a$. . . . . 8
5 Typical topographic distortions introduced by the side-looking geometry. . . . . 9
6 TanDEM-X Mission (courtesy: DLR). . . . . 10
7 Input-output relationship for the midrise (left) and midtread (right) uniform quantizers. 12
8 Input-output relationship for the Max-Lloyd optimum quantizer. . . . . 14
9 Normal distribution. . . . . 15
10 Typical training evolution scenario. . . . . 19
11 Structure of a perceptron. . . . . 20
12 Structure of a MLP architecture. . . . . 20
13 Representation of a 2D matrix convolution. . . . . 21
14 Representation of a 2D CNN convolution. . . . . 22
15 Example of error backpropagation for a perceptron with two inputs. . . . . 23
16 Basic residual learning block. . . . . 24
17 Structure of the DnCNN architecture proposed in [22]. . . . . 25

18 Block diagram of a typical spacerborne remote sensing pipeline, from acquisition to distribution. . . . . 27
19 Distribution of the complex Gaussian signal used for performance evaluation of the different quantization techniques. . . . . 28
20 Signal-to-quantization-noise ratio (SQNR) of a uniformly quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bits-per-sample. . . 29
21 Phase error standard deviation ($\sigma_{\Delta\phi}$) of a uniformly quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates. . . . . 29
22 Phase error magnitude ($|\Delta\phi|$) of a uniformly quantized, normally distributed signal for different input signal magnitudes. . . . . 30
23 Signal-to-quantization-noise ratio (SQNR) of a BAQ quantized, normally distributed signal at different signal-to-clipping ratios $\gamma_{clip}$ and for different bit-rates. . . . . 31
24 Phase error standard deviation ($\sigma_{\Delta\phi}$) of a BAQ quantized, normally distributed signal at different signal-to-clipping ratios $\gamma_{clip}$ and for different bit-rates. . . . . 32
25 Signal-to-quantization-noise ratio (SQNR) of a Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates. . 33
26 Phase error standard deviation ($\sigma_{\Delta\phi}$) of a Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates. . . . . 33
27 Signal-to-quantization-noise ratio (SQNR) comparison of a uniformly, BAQ and Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates. . . . . 34

28 Phase error standard deviation ($\sigma_{\Delta\phi}$) comparison of a uniformly, BAQ and Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates. . . . . 35
29 Intensity images of the utilized TanDEM-X acquisitions, represented in dB scale. . . . 37
30 Absolute error scatterplots compared to locally computed statistics. . . . . 40

31 Training-set generation pipeline. . . . . 45
32 Training strategy for the proposed denoiser. The CNN generates its prediction from the noisy input image $y$. The resulting denoised image $\hat{x}$ is passed through a despeckling filter $\mathcal{D}(\cdot)$. The resulting image $\hat{x}_{\mathcal{D}}$ is compared to the despeckled target image $x_{\mathcal{D}}$ to compute the loss function $\mathcal{L}(\cdot)$. . . . . 46
33 Average patch error composition for the combined dataset. The contribution of each individual acquisition is plotted as a colored dotted line. . . . . 47
34 Training-set composition after using a homogeneous (left) and heterogeneous sampling strategy (right). The contribution of each individual acquisition to the final distributions is plotted as a dotted line. . . . . 47
35 Training metric (in blue) and validation metric (in orange) evolutions for the combined dataset. The training loss function is represented in the first row, while the SQNR metric has been plotted in the second one. The first row represents the evolution using a homogeneously-sampled dataset, while the right one represents that resulting from a heterogeneously-sampled one. . . . . 48
36 Boxplots of the performance in terms of SQNR for the combined training-set. . . . . 49
37 Training loss evolution (in blue) and validation loss evolution (in orange) for the individual datasets. . . . . 51
38 SQNR boxplots for the individual test-sets. . . . . 52
39 Average performance gain versus average quantization error magnitude, using both homogeneously- and a heterogeneously-sampled training-sets. . . . . 53
40 Training loss evolution (in blue) and validation loss evolution (in orange) for the networks associated to the individual datasets. . . . . 55
41 Boxplots of the SQNR computed on the test-set. . . . . 56
42 Scatterplots of the SQNR of the input and the denoised samples in the test-set. . . . 57
43 Average performance gain using one and two input features. . . . . 58

# List of Tables


1 TerraSAR-X and TanDEM-X System Parameters [10]. . . . . 11
2 BAQ parameters for different compression rate as used for the TanDEM-X mission. . . 14
3 BAQ parameter table extended by the missing 8:5 bit and 8:7 bit entries in the TanDEM-X mission table 2. . . . . 31
4 Max-Lloyd reconstruction values as obtained by the iterative algorithm. . . . . 32
5 TanDEM-X dataset acquisition parameters. . . . . 38
6 SAR focusing correlation window for the different TanDEM-X acquisitions. . . . . 38
7 Signal performance correlation with locally computed statistics. . . . . 39
8 Sample-count comparison for the combined datasets using homogeneous and heterogeneous sampling strategies . . . . . 48
9 Average performance gains for the combined training-set. . . . . 50
10 Sample-count comparison for the individual datasets using homogeneous and heterogeneous sampling strategies . . . . . 50
11 Average performance gain for the individual training-sets. . . . . 53
12 Sample-count comparison for the individual datasets using 1 and 2 input features. . . 54
13 Average performance gains for the individual training-sets, using one and two features respectively. . . . . 54

# Contents


**1 Introduction** **1**

**2 Theoretical Background** **3**
2.1 Synthetic Aperture Radar (SAR) Principle . . . 3
2.1.1 Acquisition Geometry . . . 4
2.1.2 Image Formation . . . 6
2.1.3 Radar Parameters . . . 8
2.1.4 The TanDEM-X Mission . . . 10
2.2 SAR Raw Data Quantization . . . 11
2.2.1 Quantization Theory Fundamentals . . . 11
2.2.2 Uniform Quantizer . . . 11
2.2.3 Max-Lloyd Optimum Quantizer . . . 13
2.2.4 Block Adaptive Quantizer (BAQ) in SAR systems . . . 14
2.3 SAR Quantization Noise and Performance Parameters . . . 16
2.3.1 Granular and Clipping Errors . . . 16
2.3.2 Low Amplitude Suppression . . . 17
2.3.3 Low Scatter Suppression . . . 17
2.3.4 Performance Metrics . . . 17
2.4 Deep Learning-Based Denoising . . . 18
2.4.1 Machine Learning Fundamentals . . . 19
2.4.2 Convolutional Neural Networks . . . 21
2.4.3 Artificial Neural Network Training . . . 23
2.4.4 Deep Residual Learning . . . 24
2.4.5 Deep Neural Networks for Image Denoising . . . 24

**3 Analysis of Quantization Errors in Simulated and Real SAR Data** **27**
3.1 State-of-the-Art Quantization Methods for SAR . . . 27
3.2 Benchmarks of Quantization Methods on Simulated Data . . . 28
3.2.1 Simulated Data . . . 28
3.2.2 Uniform ADC Performance . . . 28
3.2.3 BAQ Performance . . . 31
3.2.4 Max-Lloyd Quantizer Performance . . . 34
3.2.5 Quantizer Performance Comparison . . . 34
3.3 Preliminary Analysis on Real SAR Data for Data-Rate Optimization . . . 36
3.3.1 Concept . . . 36
3.3.2 Used Dataset . . . 36
3.3.3 Experiments and Results . . . 37

**4 Proposed CNN Denoising** **43**
4.1 Concept . . . 43
4.2 Quantization Noise Model . . . 43
4.3 Proposed CNN Architecture . . . 44
4.4 Training Set Generation . . . 44

4.5 Training Strategy . . . . . 46
4.6 Experiments and Results . . . . . 47
4.6.1 Combined Dataset . . . . . 48
4.6.2 Individual Datasets . . . . . 50
4.6.3 Footprint Feature . . . . . 54

**5 Conclusion and Outlook** **59**

**Bibliography** **61**

# 1 Introduction

A Synthetic Aperture Radar (SAR) is a radar technology which uses radio waves for two-dimensional imaging purposes. Thanks to its working principles, radar-imaging is of great interest for remote sensing applications, mainly because of the possibility to operate in absence of daylight and regardless of weather conditions. The trends in SAR technology currently see increasing amounts of on-board data being generated by SAR systems, as multiple polarizations, larger bandwidths, and wider footprints are being employed. Spaceborne missions are inherently limited in the amount of acquirable data by the available on-board memory, the connection time between sensor and ground stations, and the downlink capabilities of the associated infrastructure. On-board raw data quantization thus represents a significant trade-off problem for spaceborne SAR systems, since the number of bits used for the digitization process affects both the generated data volume to be stored on-board and transmitted back to ground, as well as the quality of the generated imagery.

Different solutions have been proposed over the last years, most of which directly act on the quantization process, aiming at offering better, more efficient compression algorithms. The most commonly used quantization algorithm in on-going Earth observation SAR mission is the so-called block-adaptive quantization (BAQ) scheme. Introduced in 1989 for NASA's Magellan spacecraft [11], the BAQ algorithm, either in its original or in one of its enhanced versions, still represents the most used on-board compression architectures, being adopted by DLR's TerraSAR-X and TanDEM-X satellites [10], ASI's COSMO-SkyMed constellation [1], and ESA's Sentinel-1 constellation [2], just to name a few. More sophisticated alternatives have been proposed over the last three decades [4][16][2], but most of these trade superior compression performance for increased system complexity or computational burden (or both). On the other hand, many of today's operational constellations are past their intended life cycles, and need therefore to spare hardware consumption in the attempt to continue providing the scientific and commercial communities with valuable data. An example is the TanDEM-X mission, outliving its intended mission duration since 2016. For these existing missions, it is hard to justify significant investments to upgrade the data volume management capabilities. The master thesis work presented here fits within this context. It aims at studying the on-board quantization process and its effect on the resulting products, and at exploring the potentials of two, novel data volume reduction strategies to overcome the aforementioned limitations.

The thesis is structured as follows. Section 2 offers a theoretical overview of the topics under examination. Subsection 2.1 recalls the working principles of SAR systems. Subsection 2.2 introduces the major aspects of the quantization process, followed by a brief overview of most commonly used quantization strategies. Subsection 2.3 contextualizes the quantization mechanism in the SAR acquisition process, going over the major quantization-induced noise components and the metrics used to quantify their impact. Subsection 2.4 introduces the concepts of machine- and deep-learning, with particular emphasis on deep residual learning architectures. Section 3 covers the first half of the thesis' work, consisting of benchmarks of the most used quantization techniques, and then tests the predictability of the quantization error associated to the acquisitions of specific test-areas. Section 4 explores the potentials of deep residual learning techniques for on-ground SAR quantization denoising. The final section 5 presents the conclusions and future developments conceived from the experimental results.

# 2 Theoretical Background

## 2.1 Synthetic Aperture Radar (SAR) Principle

A Synthetic Aperture RADAR (SAR) is an airborne or spaceborne ground-mapping technology. The range of utilized radio waves renders SAR systems insensitive to cloud-coverage, allowing them to operate in any weather condition and independently from cloud coverage. Furthermore, a SAR is an active imaging system which allows to operate regardless of solar illumination. These two key aspects render SAR systems extremely flexible when compared to optical sensors, which require sunlight and a clear sky to image the area of interest. A radar system works by intermittently emitting high power, electromagnetic pulses and receiving their backscattered echoes, once every pulse repetition interval (PRI), where

$$PRI = \frac{1}{PRF} \tag{1}$$

being PRF the pulse repetition frequency. When trying to determine the position of a target considering a top-down acquisition geometry, the backscattered signal received at a certain time instant $\tau_0$ can be associated to all points on the ground sitting on the circonference defined by

$$x^2 + y^2 = \left(\frac{c_0 \cdot \tau_0}{2}\right)^2 - h_s^2 \tag{2}$$

where $x$ and $y$ are the ground plane coordinates with respect to the origin defined as the projection of the SAR sensor on the plane, $c_0$ is the speed of light in a vacuum, and $h_s$ is the height of the SAR sensor compared to the ground plane. Since the acquisition platform needs to move to keep its altitude and trajectory, the resulting Doppler shift of the points on ground can be exploited to narrow down the solution space. The Doppler analysis results in a pseudo-hyperbola of possible solutions. Intersecting this solution space with the previously defined circular one, allows one to identify four possible candidate locations. Each point is located in one of the four quarters defined by flight-direction and the ground-direction perpendicular to it. This gives rise to a left/right ambiguity and a forward/backward ambiguity, the latter of which is resolved by simply taking into consideration the sign of the Doppler shift.

The basic working principle of SAR is based on that of the Side-Looking Airborne Radar (SLAR), thus the remaining left/right ambiguity is solved by SAR systems utilizing a side-looking imaging geometry with respect to the forward motion of the acquisition platform. The resulting flight direction is denoted as azimuth direction, while the line-of-sight one as slant-range direction. It is possible to make a distinction between a monostatic radar, when the receiver antenna coincides with the transmit antenna, and a bi- or muti-static radar, when these do not coincide. Both airborne and spaceborne platforms offer advantages and drawbacks in the imaging process, allowing for application-dependent trade-offs. Satellites are typically more stable and guarantee a longer mission lifetime, whereas airborne platforms are more flexible, less expensive and potentially allow to achieve higher spatial resolutions. The azimuth resolution $\delta_a$ consists of the smallest detectable separation between two point-targets along the azimuth direction. In SLAR systems, the azimuth resolution is limited by the antenna size $L_a$, since it is inversely proportional to the azimuth antenna beamwidth $\theta_a$, where

$$\theta_a = \frac{\lambda}{L_a} \tag{3}$$

and $\lambda$ is the wavelength. Thus the azimuth resolution $\delta_a$ also depends on the range-distance $r_0$ between

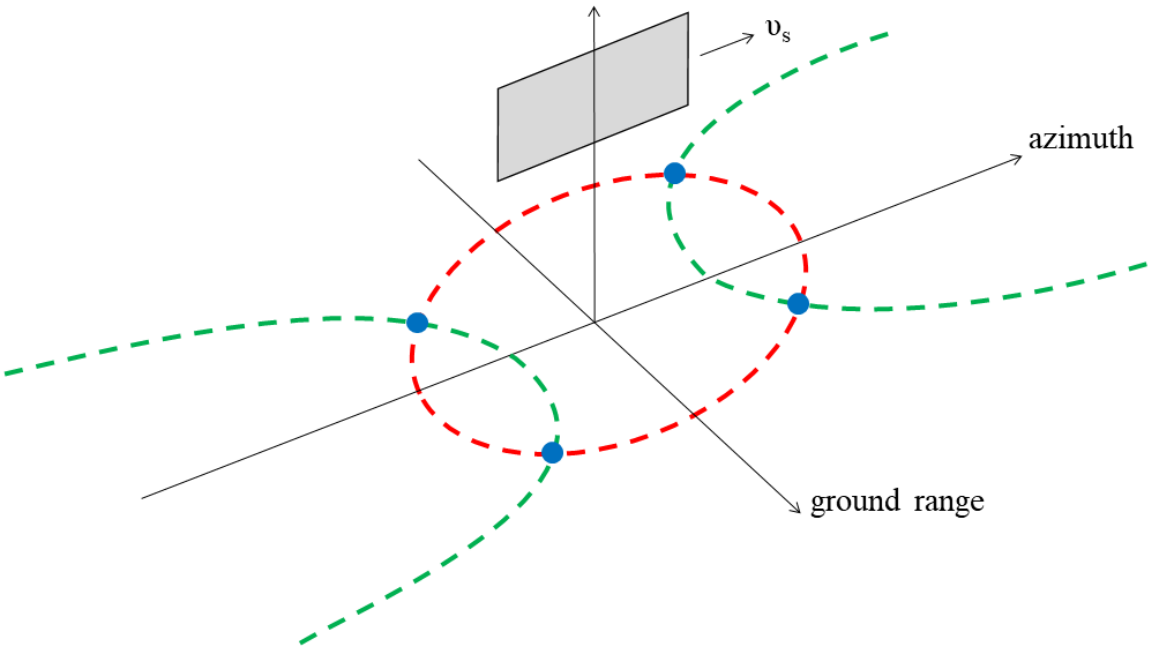



Figure 1: On the red circle lie all the possible target locations associated to the echo received at time $\tau_0$. The green pseudo-hyperbola identifies the possible target locations given the Doppler shift observed by the platform moving at speed $v_s$. The intersection identifies the remaining forward/backward and left/right ambiguities.

the scene and the acquisition platform, resulting in

$$\delta_a = \theta_a \cdot R_0 = \frac{\lambda}{L_a} R_0 \tag{4}$$

For example, a system with a 5 m antenna operating in the X band ($\lambda \approx 3$ cm) from low Earth orbit ($h_s \approx 500$ km) is going to have an azimuth resolution of 3 km. An order of magnitude increase in azimuth resolution requires a similar increase in the linear dimensions of the antenna, thus technological and mechanical constraints severely hinder the attainable performance. Similarly, changes in the utilized wavelength also follow physical and application-dependent limitations. SAR systems overcome this limitation by taking advantage of the movement of the acquisition platform along the azimuth direction to emulate an antenna with a larger, "synthetic" aperture. Due to the movement of the sensor, a scatterer on the ground stays in the antenna footprint for a given time interval, referred to as integration time, which depends on the movement speed of the platform, the azimuth antenna beamwidth, and the distance between the sensor and the acquisition scene. By appropriately combining the received backscattered signals, it is therefore possible to significantly increase the effective aperture of the antenna. The resulting azimuth resolution $\delta_a$ becomes equal to half the azimuth antenna length $L_a$

$$\delta_a = \frac{L_a}{2} \tag{5}$$

and is notably independent from the range distance $R_0$.

### 2.1.1 Acquisition Geometry

In its simplified implementation, a monostatic SAR system consists of a microwave transmitter and receiver, placed on an airplane or a satellite platform, moving at a given speed $v_s$, at a height $h_s$, along the azimuth direction. The slant-range $R_0$ is perpendicular to the azimuth direction and represents the direction toward which the system is looking, while the ground-range is the projection of the slant-range on the ground.

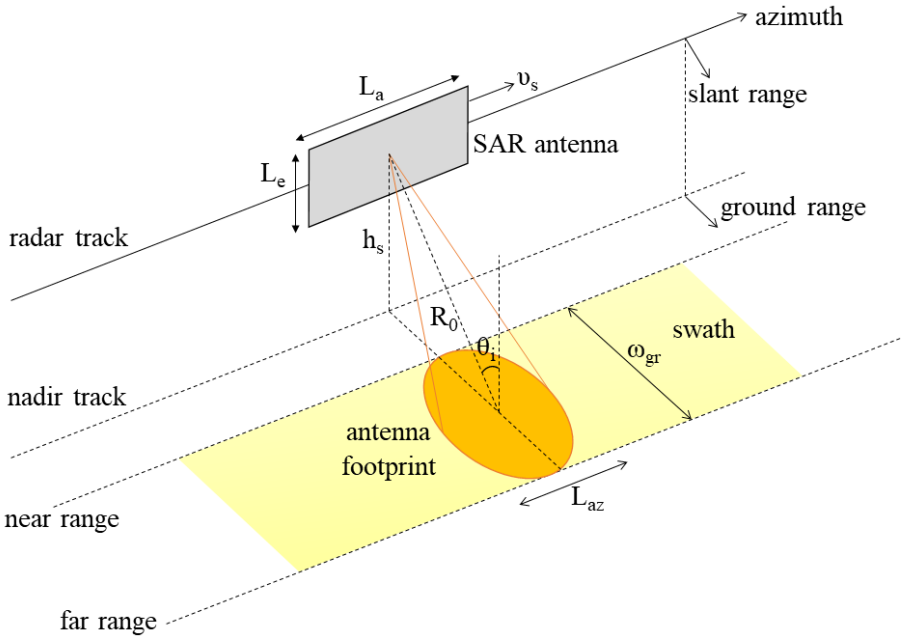



Figure 2: Typical SAR acquisition geometry.

The swath width $w_{gr}$ represents the ground-range extent of the radar scene

$$w_{gr} = \frac{\lambda \cdot R_0}{L_e \cdot \cos\theta_i} \tag{6}$$

where $L_e$ is the antenna length in elevation and $\theta_i$ is the incidence angle, defined as the angle between the slant-range direction and the normal to the surface. The swath length depends on the data take duration. A SAR is a coherent system since both the amplitude and the phase of the received echoes are recorded. These quantities are affected by the physical and dielectric properties of the targets on the ground, such as their roughness and permittivity, and by the specifications of the SAR system, for instance the wavelength of the transmitted signal, its polarisation, the antenna pattern, and the acquisition geometry. Both the transmitted and the received signals are weighted by the antenna pattern which, for a monostatic configuration, coincides for both transmission and reception. In the case of rectangular apertures, the antenna pattern gain $G$ can be approximated as

$$G\left(\psi_a, \psi_e\right) = \mathrm{sinc}^2\left(\frac{L_a}{\lambda}\psi_a\right) \cdot \mathrm{sinc}^2\left(\frac{L_e}{\lambda}\psi_e\right) \tag{7}$$

where $\psi_a$ and $\psi_e$ are the off-boresight angles in azimuth and elevation, respectively . The *sinc* functions are squared to account for the fact that the antenna pattern gain must be considered for both transmission and reception. In radar systems, the energy of the antenna pattern is typically considered within the half power beam width (HPBW), which in case of a *sinc*-pattern is approximated by

$$\psi_{a,3dB} \approx 0.886 \cdot \frac{\lambda}{L_a} \tag{8}$$

$$\psi_{e,3dB} \approx 0.886 \cdot \frac{\lambda}{L_e} \tag{9}$$

At a certain azimuth, the distance between the radar moving at constant velocity $v_s$ and the point on

the ground is found according to Pythagoras' theorem

$$R(a) = \sqrt{R_0^2 + a^2} \approx R_0 + \frac{a^2}{2 \cdot R_0} \tag{10}$$

Since $a = v_s \cdot t$, the range distance can be rewritten in function of time as

$$R(t) \approx R_0 + \frac{(v_s \cdot t)^2}{2 \cdot R_0}. \tag{11}$$

### 2.1.2 Image Formation

In a SAR system, each of the received echoes is a superposition of the responses of all elementary scatterers present within the antenna beam. All the acquired echoes are arranged side-by-side in a bi-dimensional array, in which each data cell is identified by the time delay and the pulse number. As a consequence of the central limit theorem applied to the large number of individual scatterers contributing to each received echo, the in-phase and quadrature components of the signal can be modelled as zero-mean, stationary and independent processes. This data arrangement is typically referred to as raw data. Each received echo is sampled (i.e. discretization in time) and then separately quantized (i.e. discretization in amplitude) for the in-phase and quadrature components with a predefined number of bits-per-sample (bps) $N_b$ by the on-board analog-to-digital converter (ADC). In a second moment, the resulting digitized raw data is downlinked to the ground. Since each data unit belonging to the 2D raw-data matrix consists of a superposition of the responses of multiple scatterers, the resulting image is said to be defocused and thus further processing is required. This accordingly named focusing step is essential to the formation of the final 2D image, and consists in the convolution of the raw data matrix with the system's reference function in range and azimuth.

#### Range Focusing

Let $s(t)$ be the waveform transmitted by the SAR sensor at a certain time $t$. The signal $x(t)$ backscattered from a point-like object, will be an attenuated, delayed and phase-shifted version of $s(t)$ with an additive white Gaussian noise (AWGN) contribution $\eta(t)$

$$x(t) = s\left(t - \frac{2R}{c_0}\right) \cdot exp\left(j\frac{4\pi R}{\lambda}\right) + \eta(t) \tag{12}$$

Under this assumption, the optimal linear time invariant (LTI) filter to maximize the signal-to-noise ratio (SNR) is a so-called matched filter. Indeed, in radar systems range focusing is implemented by convolving the received signal with the conjugated, time-reversed replica of the transmitted signal

$$\chi(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(\tau) \cdot s^*\left(-(t-\tau)\right) \cdot d\tau \tag{13}$$

In order for two different scatterers to be separately detected, their echoes must be separated by a large enough time delay. Under the simplistic assumption of a rectangular waveform of duration $\tau_p$ transmitted by the sensor, it is possible to show that the geometry of the problem limits the spatial resolution in slant-range to be

$$\delta_{sr,rect} = \frac{c_0 \tau_p}{2} \approx \frac{c_0}{2B_\tau} \tag{14}$$

where $c_0$ is the speed of light in a vacuum and the relationship $\tau = \frac{1}{b_\tau}$ has been exploited. This implies that, in order to improve the slant-range resolution, the transmitted pulse-length must be shortened, this in turn requires increasing the transmission power to maintain the same SNR. Technical constraints overall limit the maximum amount of radiated power and thus the resolution/SNR trade-off. To overcome this limitation, radar systems usually utilize a complex linear frequency modulated

(LFM) pulsed waveform known as chirp

$$s(t) = exp\left[j\left(\theta_0 + 2\pi f_c t + \pi k_r t^2\right)\right] \cdot rect\left(\frac{t}{\tau_p}\right) \tag{15}$$

where $f_c$ is the carrier frequency, $\theta_0$ is the initial phase, and $k_r$ is the chirp rate. Note that the resulting waveform duration is limited in time by the window function to be equal to $\tau_p$. The resulting chirp bandwidth can be computed as

$$B_{rg} = \tau_p \left|k_r\right| \tag{16}$$

which implies that for chirp signals, the resolution improves as longer waveforms are transmitted, thus also increasing the transmitted energy and the achievable SNR at the receiver.

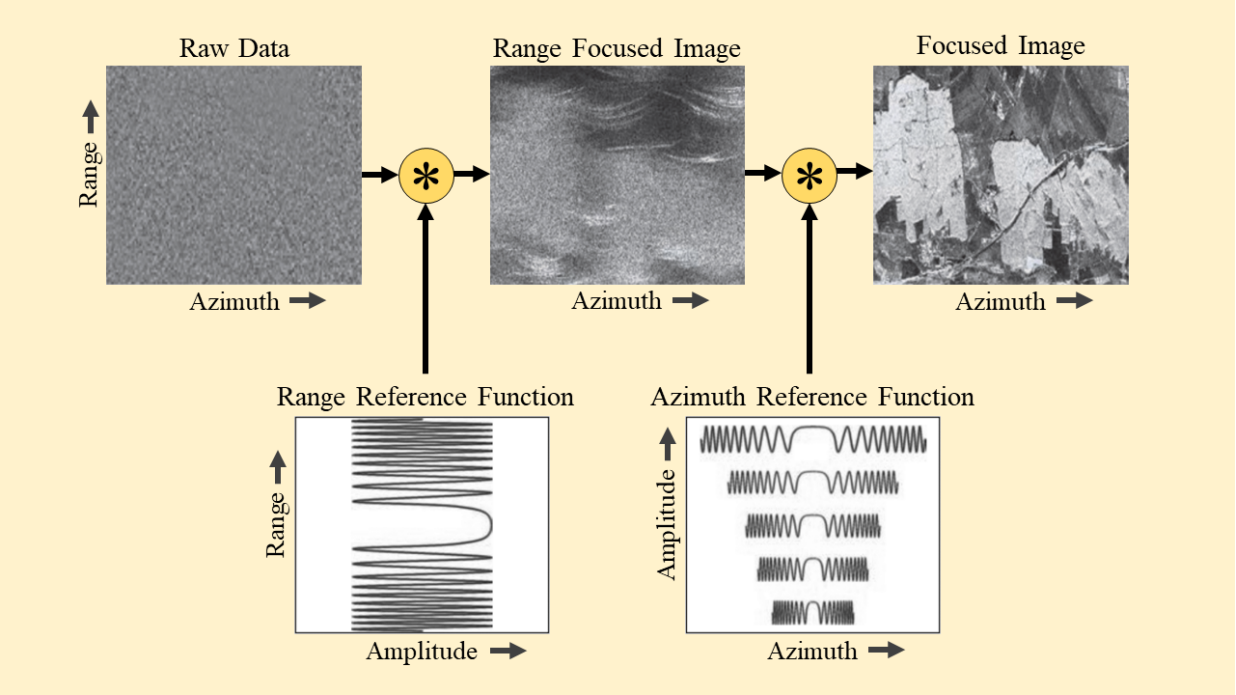



Figure 3: SAR focusing process.

### Azimuth Focusing

Similarly to range focusing, azimuth focusing consists in convolving the received echoes along the azimuth lines with the expected response from a point target on the ground

$$s(t) = A\sqrt{\sigma_0} exp\left[j\left(\theta_s - \frac{4\pi}{\lambda}R(t)\right)\right] \cdot rect\left(\frac{v_s t}{L_S}\right) \tag{17}$$

where $\sigma_0$ is the scattering coefficient, $\theta_s$ is the scattering phase and $R(t)$ is the time-varying distance between the sensor and the point-like target. Note that the function is limited in space by the sensor azimuth position $a = v_s t$ and the synthetic antenna azimuth length $L_s$, which is equal to the synthetic aperture $L_{sa}$. The attainable azimuth resolution for SAR system has been defined in (5) Intuitively, this result can be explained by the fact that a longer distance allows for a larger synthetic antenna, which in turn is more directive: the principle of synthetic aperture directly compensates the effects of the sensor-to-target distance.

Further constraints regarding the sampling process need to be taken into consideration. Similarly to the definition of range bandwidth, the doppler bandwidth $B_D$ can defined as the reciprocal of the

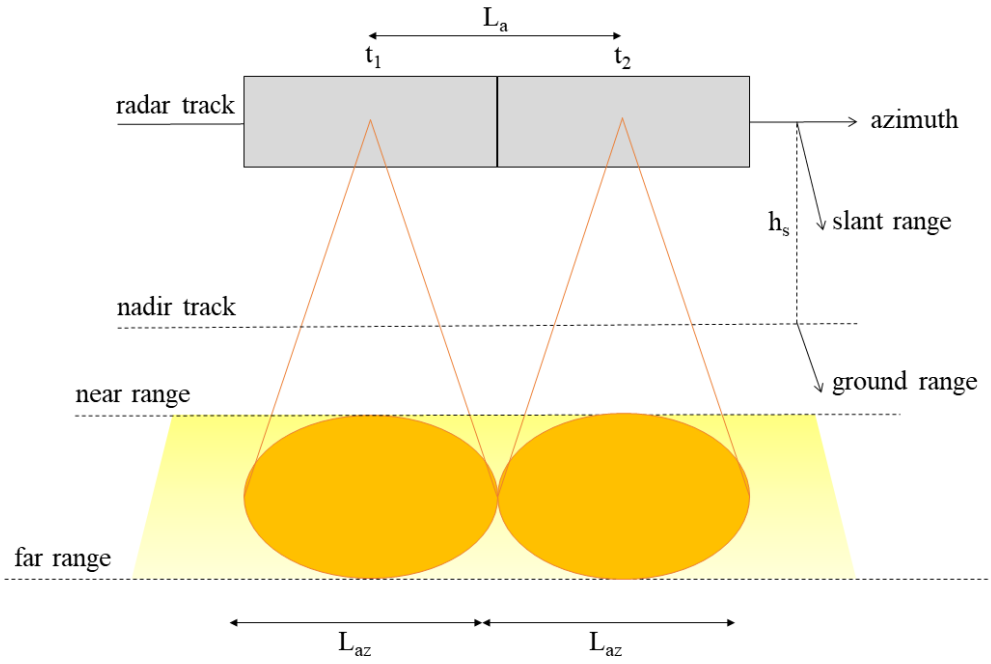



Figure 4: SAR system at instants $t_1$ and $t_2$. The resulting synthetic aperture is equal to halt the antenna aperture in azimuth $L_a$.

time required by the sensor to cover one resolution cell

$$\delta_a = \frac{v_s}{B_D} \tag{18}$$

This result implies that a trade-off between azimuth resolution and swath width has to be made.

### 2.1.3 Radar Parameters

#### Radar Backscatterer and Speckle

From the point of view of the SAR system, scatterers can be split into two main categories: point scatterers and distributed scatterers. A point target consists of a single dominant scatterer within the resolution cell. The power of the signal backscattered by such a target can be described by the point target radar equation, which for a monostatic configuration, can be expressed as follows

$$P_q^{rx} = \frac{P_q^{tx} G_p G_q \lambda^2}{(4\pi)^3 R^4} \sigma_{pq} \tag{19}$$

where $P_q^{tx}$ is the power of the $q$-polarized transmitted signal, $G_p$ and $G_q$ are the transmit- and receive-antenna gains, and $\sigma_{pq}$ the radar cross section. The radar cross section quantifies the fraction of power density reflected to the sensor with respect to the overall power density, incident on the target; for a point target it is defined as

$$\sigma_{pq} = \lim_{R \to \infty} \left\{ 4\pi R^2 \frac{S_q^{rx}}{S_q^{tx}} \right\} \tag{20}$$

where the limit denotes the far-field approximation. Distributed targets consist of the superposition of multiple, non-dominant scatterers within the same resolution cell, thus they can be modelled as a normal distribution following the central limit theorem. For distributed targets, the total received power is obtained by coherently summing the contributions defined by the radar equation for all

scatterers in the illuminated area

$$P_q^{rx} = \frac{P_q^{tx} G_p G_q \lambda^2}{(4\pi)^3} \sum_i \left\{ \frac{\sigma_{pq}}{R_i^4} \right\}. \tag{21}$$

This superposition of contributions of randomly distributed scatterers within the same resolution cells, causes spatial variability in the pixel intensities. This emerging phenomenon takes the name of speckle and is typically modelled as a multiplicative process, meaning that its impact scales with the intensity of the signal. Contrary to other sources of disturbance like system noise, it is thus not possible to suppress it by raising the transmission power. In order to reduce the spatial variability of the phenomenon, low-pass filtering techniques such as multi-looking are used. It consists of a non-coherent average of $N$ neighbouring pixels, reducing the overall resolution by a factor of $N$ and the backscatter standard deviation by a factor of $\sqrt{N}$. Thus, the larger the number of looks is, the higher the radiometric accuracy and the lower the spatial resolution are. Multi-looking can be implemented working in different domains:

- Spatial Domain: a local averaging process is applied to the pixels using a sliding window;
- Time Domain: divide the synthetic aperture into multiple sub-segments, each leading to a separately processed, lower resolution image;
- Frequency Domain: similarly to the time domain approach, the doppler spectrum is divided into sub-bands, each separately processed.

### Geometric Distortions

The side-looking acquisition geometry used by SAR systems, introduces geometrical distortions as a consequence of trying to map a 3D surface onto a 2D plane. Three distinct phenomena may occur, depending on the relative position of the sensor to the elevation differences in the scene:

- Foreshortening: a point is wrongfully seen as being closer to the sensor than it effectively is, thus affecting the relative distance between points as seen from a top-down perspective;
- Layover: the relative distance between two points on the ground is inverted in the slant-range perspective;
- Shadow: the topography of the scene causes some of its parts to obscure other ones.

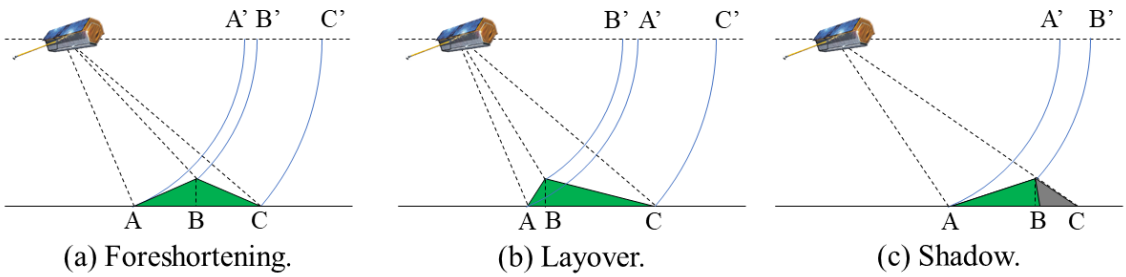



Figure 5: Typical topographic distortions introduced by the side-looking geometry.

### Noise Equivalent Sigma Zero

The noise equivalent sigma zero (NESZ) represents the backscattering coefficient required to obtain a unitary signal-to-noise ration, $SNR = 1$. It can be interpreted as the noise floor of the SAR system,

thus it is a measure of its sensitivity, with lower NESZ values being better. It can be derived from the radar equation as

$$NESZ_{pq} = \frac{4^4\pi^3 R^3 v_s \sin\eta k_B T_{rx} B_{rg} F_N L_{tot}}{P_p^{tx} G_p G_q \lambda^3 c_0 \tau_p} \cdot PRI \tag{22}$$

where $k_B$ is the Boltzmann constant, $T_{rx}$ is the receiver temperature, $F_N$ is the noise figure and $L_{tot}$ is the sum of all losses.

### 2.1.4 The TanDEM-X Mission

Developed under a public-private partneship between the German Aerospace Center (DLR) and Airbus Defense & Space (ADC) [10], TanDEM-X is a spaceborne bistatic X band SAR system. It is comprised of two separate spacecrafts, the TerraSAR-X satellite (TSX) launched in 2007 and its twin companion satellite TanDEM-x (TDX). Some of the main TerraSAR-x and TanDEM-X system parameters are presented in Table 1. The primary objective of this mission was the creation of a global, high-precision

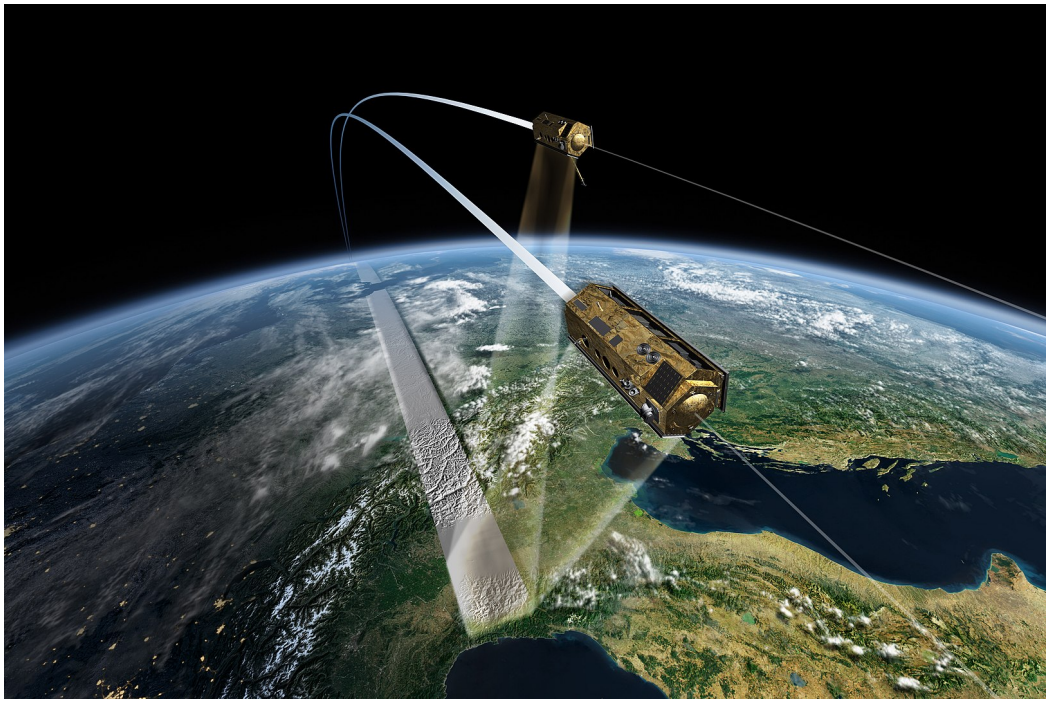


Figure 6: TanDEM-X Mission (courtesy: DLR).

digital elevation model (DEM) [10], sucessfully completed in 2016. Since the official mission start in December 2010, the constellation has been acquiring interferometric SAR images in stripmap, single HH polarization mode, with a typical azimuth and range resolution of 3 m. For the generation of the final DEM, a boxcar multilooking filter is used for phase filtering and coherence estimation, lowering the final spatial resolution of the product to 12 m. The satellites fly in sun-synchronous orbit, combined into the so-called HELIX satellite formation [10], consisting of two vertically and horizontally displaced orbits. This orbit configuration was chosen for flexible selection of the baselines. Typically, the twin radar system operates using a bistatic acquisition scheme, with both satellites simultaneously recording the return signal transmitted by one of the two systems. When required, they can also operate in pursuit monostatic mode, in which both platforms operate independently.

| Parameter | Value |
|---|---|
| Orbit Height | 511.5 km |
| Platform Velocity | 76000 m/s |
| Carrier Frequency | 9.65 GHz |
| Chirp Bandwidth | 100 MHz |
| Sampling Frequency | 110 MHz |
| Antenna Length | 4.8 m |
| Antenna Width | 0.7 m |
| Peak TX Power | 2260 W |
| Duty Cycle | 18.0 % |

Table 1: TerraSAR-X and TanDEM-X System Parameters [10].

## 2.2 SAR Raw Data Quantization

The analog signals received by the SAR sensor need to be digitized in order for them to be stored on-board and later on sent to the ground to be processed. This analog-to-digital conversion (ADC) consists of a time discretization step called sampling, followed by an amplitude discretization step called quantization. The sampling-rate and the number of bits-per-sample (bps) constitute a typical trade-off scenario as they determine the final quality of the digitized signal and the overall memory requirements. The typical SAR scenario requires the usage of lossy compression techniques to overcome the capacity limits of the on-board memory. These compression techniques are usually not computationally high demanding as to limit excessive usage of platform resources.

### 2.2.1 Quantization Theory Fundamentals

Quantization essentially consists in mapping the dynamic of the discrete-time (DT) signal $x[n]$ onto the nearest value $x_q[n]$ out of a set of predetermined possible outputs, covering the signal's dynamic range. The quantization process thus introduces a certain amount of error depending on the number of possible representations $M = 2^{N_b}$, where $N_b$ is the number of bits-per-sample. The set of $M$ discrete reconstruction values implicitly defines the dynamic of the quantizer. Input values outside of this interval are clipped before being quantized, de facto discretizing the dynamic range of an otherwise unbounded input signal. The error $q$ introduced by the quantization process, typically referred to as quantization noise, can thus be seen as consisting of two different phenomena:

- Granular error: consisting of the distortion caused by the round-off operation for values within the supported range. It can be modelled as an additive, deterministic, randomly distributed noise source
$$q[n] = x[n] - x_q[n],\ with\ x(t) \in [d_i; d_{i+1}] \tag{23}$$
where $[d_i; d_{i+1}]$ represents the $i^{th}$ decision interval.
- Overload/Clipping Error: consisting in the distortion introduced by the clipping operation onto input values lying outside of the supported range. This is typically a non-linear noise effect.

### 2.2.2 Uniform Quantizer

The input-output relation function of a quantizer is characterized by being a non-linear, staircase-like function, relating the $M$ input intervals to the corresponding $M$ output representations. As mentioned before, the input values outside the $[-V_{clip}; +V_{clip}]$ range are clipped and associated to the corresponding output boundary values $\pm r_{clip}$. If all the input intervals cover a uniformly-spaced step-size $\Delta$, then we talk about a uniform quantizer, otherwise the quantizer is said to be non-uniform. Depending on the arrangement of the decision intervals, it is then possible to define either a midriser or a midtread quantizer (figure 7). The midriser type is characterized by the origin point being a threshold value between two adjacent decision intervals, thus not possessing the zero-value as a possible output reconstruction; for a uniform quantizer, the input-output relation can be expressed as

$$x_q[n] = \Delta \cdot \left( \left\lfloor \frac{x[n]}{\Delta} \right\rfloor + \frac{1}{2} \right),\ if\ x \in [-V_{clip}; +V_{clip}] \tag{24}$$

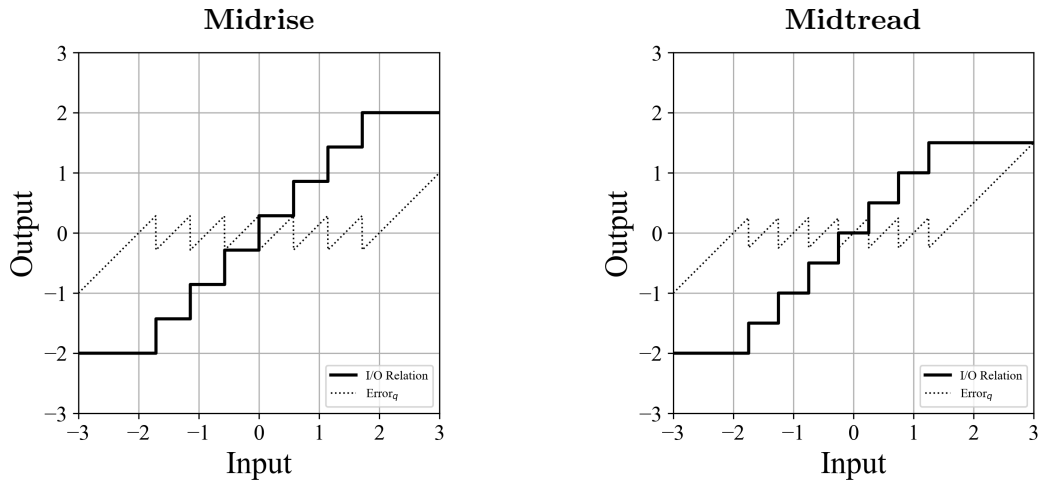



Figure 7: Input-output relationship for the midrise (left) and midtread (right) uniform quantizers.

where $\lfloor \cdot \rfloor$ represents the so-called "floor" function, returning the greatest integer less than or equal to its argument. A midtread quantizer possesses a zero-valued reconstruction output $x_q[n] = 0$; its input-output relation can be expressed as

$$x_q[n] = \Delta \cdot \left\lfloor \frac{x[n]}{\Delta} + \frac{1}{2} \right\rfloor, \; if \; x \in [-V_{clip}; +V_{clip}] \tag{25}$$

The decision of using a uniform quantization scheme over a non-uniform one is pragmatically driven by the model of the input function. A uniformly distributed signal prompts the usage of a uniform quantizer, while in every other case a non-uniform quantizer can be optimized in its decision level layout, in order to minimize the overall mean squared error (MSE) of the quantization noise. In general, the step-size $\Delta$ must be chosen accordingly to the statistics of the input signal.

Let $x[n]$ be a stochastic process with zero mean, variance $\sigma_x^2$ and described by the probability density function (pdf) $p_x(x)$, that is

$$\mu_x = \mathbb{E}[x] = \int_{-\infty}^{+\infty} x \cdot p_x(x) \cdot dx = 0, \tag{26}$$

$$\sigma_x^2 = \mathbb{E}[x^2] = \int_{-\infty}^{+\infty} x^2 \cdot p_x(x) \cdot dx \tag{27}$$

where $\mathbb{E}[\cdot]$represents the expectation operator. As a consequence, the quantization error $q$ is also a zero mean stochastic process, described by its pdf-function $p_q(q)$ and variance

$$\sigma_q^2 = \mathbb{E}[q^2] = \int_{-\infty}^{+\infty} (x - x_q)^2 \cdot p_x(x) \cdot dx \tag{28}$$

where the equation has been written in function of the input signal $x[n]$. By integrating over all the $M$ decision intervals, the resulting quantization error power becomes

$$\sigma_q^2 = \mathbb{E}[q^2] = \sum_{k=1}^{M} \left\{ \int_{d_k y}^{d_{k+1}} (x - x_q)^2 \cdot p_x(x) \cdot dx \right\} \tag{29}$$

with $d_1 = -\infty$ and $d_M = +\infty$. Under the assumption of a uniform quantizer, the step size can be computed as:

$$\Delta = \frac{2V_{clip}}{2^{N_b}}. \tag{30}$$

While the number of representation levels is determined by the number $N_b$ of utilized bits-per-sample, $V_{clip}$ must be chosen to reduce the overall quantization error; larger values of $V_{clip}$ minimize the overload error but increase the granular error for a given number of bps. To this end, the input clipping value $V_{clip}$ must be as representative as possible of the dynamic range of the signal

$$|x| \leq V_{clip} \tag{31}$$

If the step size is small enough, it is reasonable to assume that the quantization error is uniformly distributed within the interval

$$p_q(q) = \begin{cases} \frac{1}{\Delta} & \text{, if } |q| < \frac{\Delta}{2} \\ 0 & \text{, otherwise} \end{cases} \tag{32}$$

Under these assumptions it is possible to express the quantization error power in closed form

$$\sigma_q^2 = \frac{\Delta^2}{12} = \sigma_x^2 2^{-2N_b} \tag{33}$$

which shows that the variance of the quantization error increases with square of the step size and exponentially decreases with the number of bits-per-sample.

### 2.2.3 Max-Lloyd Optimum Quantizer

In the case of a normally or otherwise non-uniformly distributed stochastic process, the uniform quantizer allocates the same number of resources to discretize each interval, regardless of their respective probability density functions. A basic principle of information theory consists in allocating more resources to those values which are more likely to be occurring and vice versa. Under this assumption, the optimum quantizer is the design for which a fixed number of bits $N_b$ defines the distribution of the intervals $d_i$ and their associated reconstruction levels $r_i$, so as to minimize the quantization noise power

$$\{\boldsymbol{d_{opt}}, \boldsymbol{r_{opt}}\} = \arg\min_{\{\boldsymbol{d},\boldsymbol{r}\}} \left\{\sigma_q^2(N_b)\right\} \tag{34}$$

Given the pdf of input signal $p_x(x)$, minimizing the variance of the noise is equal to find the values $\boldsymbol{d_{opt}}$ and $\boldsymbol{r_{opt}}$ for which the derivative of the variance is null

$$\frac{\partial \sigma_q^2}{\partial d_k} = 0 \;\Rightarrow\; (d_k - rk - 1)^2 = (d_k - rk)^2\,,\; if\; p_x(d_x) \neq 0,\; \forall k = 2, 3, .., M. \tag{35}$$

$$\frac{\partial \sigma_q^2}{\partial r_k} = 0 \;\Rightarrow\; \int_{d_k y}^{d_{k+1}} (x - r_k) \cdot p_x(x) \cdot dx = 0,\; \forall k = 1, 2, .., M. \tag{36}$$

where $M = 2^{N_b}$.By solving for $\boldsymbol{d_{opt}}$ and $\boldsymbol{r_{opt}}$, we obtain

$$d_{opt,k} = \frac{(r_{opt,k} + r_{opt,k+1})}{2},\; \forall k = 2, 3, .., M,\; d_{opt,1} = -\infty,\; d_{opt,M+1} = +\infty \tag{37}$$

$$r_{opt,k} = \frac{\int_{d_{opt,k}}^{d_{opt,k+1}} x \cdot p_x(x) \cdot dx}{\int_{d_{opt,k}}^{d_{opt,k+1}} p_x(x) \cdot dx},\; \forall k = 1, 2, .., M. \tag{38}$$

Thus, the optimal decision levels are the center-points of neighbouring reconstruction levels $r_{opt}$, while these are the centroids of the signal pdf in the corresponding decision interval. Such a quantizer is often referred to as Max-Lloyd quantizer and is said to be optimal as it minimizes the resulting mean-squared error (MSE). It is intuitive to note that in the case of a uniformly distributed input signal, the optimal quantizer will coincide with a simple uniform one. For a generic signal pdf, there is no analytical solution to the equations. Numerical approximations of the solutions can be computed using an iterative process.

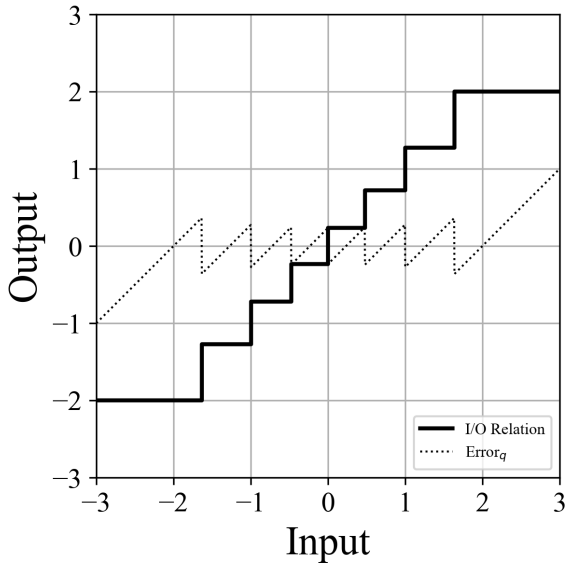



Figure 8: Input-output relationship for the Max-Lloyd optimum quantizer.

### 2.2.4 Block Adaptive Quantizer (BAQ) in SAR systems

Typical SAR echos are characterized by a large dynamic range spanning several decibels due to the presence of both low and high reflectivity targets. In order to properly digitize such signals, using a simple uniform ADC would require the usage of a very large bit-rate. However, this is in contrast with on-board memory limitations, which require the quantization bit-rate to be kept as low as possible in order to acquire useful amounts of data. To overcome this limitation, SAR systems often deploy adaptive quantizers to compress echoes digitized using high bit-rate, uniform quantizers to a more manageable data stream. One of the most widely used techniques on-board of SAR systems is the Block Adaptive Quantization (BAQ) scheme. It is a lossy data reduction technique which uses the statistics of blocks of data to set their re-quantization intervals and reconstruction levels. The size $N_{block}$ of these blocks is chosen to be large enough to allow for the correct block statistics estimation, but small enough to preserve the homogeneity of local statistics. The compression might be applied in either time or frequency domains. Time domain compression leads to higher signal-to-noise ratio, but it does not optimize the compression scheme for the spectral envelope of the signal power in the range and azimuth directions. On the other hand, compression in the frequency domain achieves better compression performance at the cost of higher system complexity. As a consequence of the central limit theorem, BAQ is usually implemented using a cartesian quantization scheme with the In-Phase and Quadrature components of the complex echo treated separately. A polar quantization scheme might also be used, even though it has been demonstrated, that no significant improvement can be achieved [17]. For example, both DLR's TerraSAR-X and TanDEM-X satellites make on-board use of

| **Compression Rate** | $C$ **Value** | $E_{max}$ **Value** | $M_{max}$ **Value** |
|---|---|---|---|
| 8:2 | 2.20374 | 24 | 1 |
| 8:3 | 5.28038 | 20 | 3 |
| 8:4 | 8.50475 | 16 | 7 |
| 8:6 | 15.25490 | 8 | 31 |

Table 2: BAQ parameters for different compression rate as used for the TanDEM-X mission.

a cartesian BAQ scheme for raw-data compression. In this particular case, the input signal is firstly clipped within the $V_{clip} = \pm 127.5mV$ range and then digitized using an 8-bit, uniform ADC. The

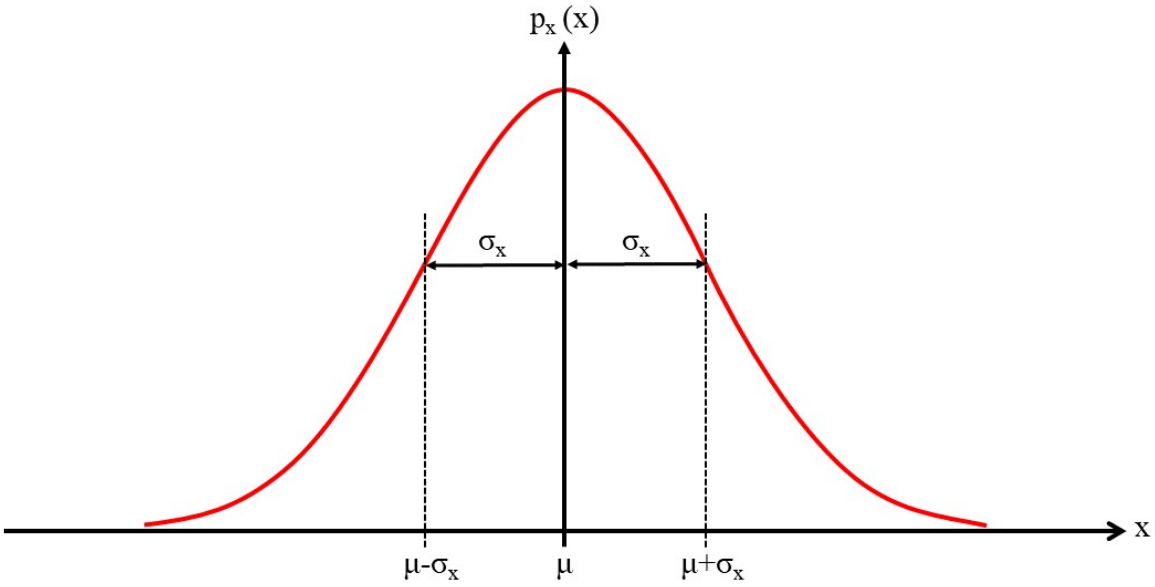



Figure 9: Normal distribution.

corresponding decimal representation is simply obtained as follows

$$I_{adc} = (-1)^s \cdot \left( \frac{1}{2} + \sum_{i=2}^{N_b-2} \{K_i 2^i\} \right) \tag{39}$$

$$Q_{adc} = (-1)^s \cdot \left( \frac{1}{2} + \sum_{i=2}^{N_b-2} \{K_i 2^i\} \right) \tag{40}$$

where the $N_b^{th}$-bit is the sign bit of in-phase and quadrature components; the resulting values belong into the $[-127.5; +127.5]$ range. For the TerraSAR-X and TanDEM-X satellites, the BAQ compression pipeline consists of the following steps:

- The output of the ADC ($I_{adc}$ and $Q_{adc}$) is divided along the range line into blocks of $N_{block} = 128$ samples. Depending on the acquisition parameters, this quantity corresponds to a range extension of about 70 to 200 m; where $[d_i; d_{i+1}]$ represents the $i^{th}$ decision interval.

- The $C$ constant corresponding to the BAQ compression rate is taken from table 2;

- The block exponent $E_1$ is computed as follows:

$$E_1 = 4 \cdot log_2 \left( 1 + \frac{1}{N_{block}} \sum_{n=1}^{N_{block}} \{|I_{adc,n}| + |Q_{adc,n}|\} \right) - C; \tag{41}$$

- The resulting exponent value $E_1$ is compared with the $E_{max}$ value corresponding to the BAQ compression-rate entry in table 2. The final exponent $E$ is computed as

$$E = min\{E_{max}, \lfloor E_1 \rfloor\} \tag{42}$$

with $min\{\cdot\}$ returning the minimum of its arguments and $\lfloor E_1 \rfloor$ returning the closest integer less than or equal to $E_1$. The resulting block exponent is coded using 5-bits;

- The computed exponent is used to compress the dynamic range of the raw data samples

$$\overline{I}_n = \frac{I_{adc,n}}{2^{\frac{E}{4}}} \tag{43}$$

$$\overline{Q}_n = \frac{Q_{adc,n}}{2^{\frac{E}{4}}} \tag{44}$$

- The mantissa value of each sample is computed by comparing the scaled samples to the reference Mmax value found in the BAQ table 2 entry for the current compression rate

$$M_I = \frac{\overline{I}}{\left|\overline{I}\right|} \cdot min\left\{\left|\overline{I}\right|, M_{max}\right\} \tag{45}$$

$$M_Q = \frac{\overline{Q}}{\left|\overline{Q}\right|} \cdot min\left\{\left|\overline{Q}\right|, M_{max}\right\} \tag{46}$$

  where $\frac{\overline{I}}{\left|\overline{I}\right|}$ and $\frac{\overline{Q}}{\left|\overline{Q}\right|}$ are the sample signs.

- The resulting mantissa is quantized in the $[-M_{max} - \frac{1}{2}; +M_{max} + \frac{1}{2}]$ range using a $N_{baq}$-bit uniform quantizer.

The resulting quantized blocks are stored on-board and later transmitted to ground. The SAR raw data values are reconstructed block-by-block by simply applying the following equations:

$$I_{baq} = (-1)^s \cdot \left(\frac{1}{2} + \sum_{i=2}^{N_{baq}-2} \left\{K_i 2^i\right\}\right) 2^{\frac{E}{4}} \tag{47}$$

$$Q_{baq} = (-1)^s \cdot \left(\frac{1}{2} + \sum_{i=2}^{N_{baq}-2} \left\{K_i 2^i\right\}\right) 2^{\frac{E}{4}}. \tag{48}$$

## 2.3 SAR Quantization Noise and Performance Parameters

In SAR systems, the ADC process is a major source of signal-dependent noise, affecting the digitized signal stored on-board. Two contributions to this phenomena have been mentioned: granular noise, which is due to a finite number of values trying to reconstruct an analog random stochastic process, and overload noise, which is caused by the clipping of the input signal exceeding the supported dynamic range. Heterogeneities in the backscatter distribution constitute a third source of signal-dependent performance degradation caused by the quantization process in SAR imaging. Such an effect is known as low-scatter suppression and is a drawback of the BAQ compression algorithm.

### 2.3.1 Granular and Clipping Errors

Let us assume that $x[n]$ is a stochastic process with zero mean and described by its pdf $p_x(x)$. Discretizing such a discrete-time signal with a finite set of reconstruction levels to choose from, entails the introduction of quantization noise in the form of a round-off error. Indeed, samples are discretized by associating them to the nearest available representation value. The resulting reconstruction error takes the name of granular error. Its value is maximum when the input sample lies on the threshold between two adjacent decision regions. The standard deviation of granular noise can be expressed as

$$\sigma_{err_g} = \sum_{i=1}^{M} \left\{ \int_{r_i - \frac{\Delta}{2}}^{r_i + \frac{\Delta}{2}} (x - r_i)^2 \cdot p_x(x) \cdot dx \right\} \tag{49}$$

where $M = 2^{N_b}$ reconstruction levels $r_i$, $\forall i = 1, .., M$ have been assumed. The quantization process requires that the input signal possesses a finite dynamic range, supported by the set of reconstruction values. By clipping the sample-values lying outside of the supported $[-V_{clip}; +V_{clip}]$ range, the dynamic

range is effectively discretized, but a clipping (or overhead) error is introduced. Its standard deviation can be computed as

$$\sigma_{err_o} = \int_{-\infty}^{r_1 - \frac{\Delta}{2}} (x - r_1)^2 \cdot p_x(x) \cdot dx + \int_{r_M + \frac{\Delta}{2}}^{+\infty} (x - r_M)^2 \cdot p_x(x) \cdot dx \tag{50}$$

The chosen quantizer dynamic has also an impact in terms of granular noise, as trying to minimize overhead errors by extending the supported dynamic range causes the available reconstruction samples to be less representative as the step-size between them grows, leading to an increase of the overall granular noise power. The overall quantization error power can thus be derived as the sum of the granular- and overload-error variances

$$\sigma_q^2 = \sigma_{err_g}^2 + \sigma_{err_o}^2. \tag{51}$$

### 2.3.2 Low Amplitude Suppression

The use of a Cartesian quantizer hides another form of error introduced by the ADC process. In terms of phase error, quantization noise is strongly dependent on the corresponding signal amplitude. The phase error is bounded to a maximum of 45° since

$$|\Delta\theta|_{max} = \frac{|\theta_{x_a,y_a} - \theta_{x_b,y_b}|}{2} \tag{52}$$

This error is highest for samples with lower amplitude, regardless of the amount of granular noise. Therefore, those areas on the ground comprising small backscatterers are going to be reconstructed with less fidelity, decreasing the final performance of phase-based techniques such as interferometry.

### 2.3.3 Low Scatter Suppression

Inhomogeneities in the distribution of backscatter in the scene contribute to another form of error during the ADC process, which is known as low scatter suppression. In particular, during the acquisition phase, if the distance between two (or more) targets is much smaller in azimuth ($d_{az}$) and range ($d_{rg}$) than the synthetic aperture of the antenna and the chirp length respectively, their responses overlap

$$d_{az} << \lambda \frac{R_0}{L_a} \tag{53}$$

$$d_{rg} << \lambda \frac{c_0 \tau_p}{2} \tag{54}$$

The BAQ algorithm adapts the dynamic of the re-quantization step to the local statistics of the signal, thus for overlapping targets with different magnitude responses the weaker signal will be heavily distorted by the compression process. The impact of such a phenomena becomes visible only after image focusing.

### 2.3.4 Performance Metrics

In order to describe a quantizer and evaluate its performance, the most relevant metrics are going to be detailed in the following.

#### Signal-to-Clipping Ratio ($\gamma_{clip}$)

It measures the ratio between the signal's standard deviation $\sigma_x$ and the clipping value $V_{clip}$ utilized during the quantization step

$$\gamma_{clip} = \frac{\sigma_x}{V_{clip}} \tag{55}$$

When considering a complex signal and considering the standard deviations of its components to be the same, we have that

$$\gamma_{clip} = \frac{\sqrt{\sigma_{x,I}^2 + \sigma_{x,Q}^2}}{V_{clip}} = \sqrt{2}\frac{\sigma_{x,I}}{V_{clip}} \tag{56}$$

The metric relates the dynamic range of the input signal to the quantizer dynamic, thus $\gamma_{clip}$ should be neither too large (i.e. dominant clipping noise), nor too small (i.e. dominant granular noise).

#### Signal-to-Quantization-Noise Ratio ($SQNR$)

It is defined as the ratio between the power of input signal to the quantizer and the quantization noise power

$$SQNR = \frac{\sigma_x^2}{\sigma_q^2} \tag{57}$$

When working with complex signals, the metric can be either re-written for its components separately, or as a sum of intensity values as follows

$$SQNR = \frac{\sum_{i=1}^{N} \left\{ |x_i|^2 \right\}}{\sum_{i=1}^{N} \left\{ |q_i|^2 \right\}} \tag{58}$$

The metric describes how much the output signal has been corrupted by the quantization noise. Assuming that the quantization error can be approximated with a uniform model, it is possible to show, by substituting (33) in (57), that for an optimum quantizer the performance in dB is given by

$$SQNR_{dB} = 10 \cdot log_2(\sigma_x^2) - 10 \cdot log_2(\sigma_x^2 2^{-2N_b}) \approx 6 \; dB/N_b \tag{59}$$

thus for each additional bit per quantized sampled, a gain of approximately 6 dB in SQNR is obtained. Such a result is a well known approximation from rate-distortion theory.

#### Phase Error ($\Delta\phi$)

For interferometric products it is of particular interest to model the impact of the quantization operation on the phase error. Such a single-point phase error is defined as the phase difference between the quantized signal and the original one

$$\Delta\phi = \arctan\left(x \cdot x_q^*\right). \tag{60}$$

#### Interferometric Coherence ($\gamma$)

The coherence product $\gamma$ quantifies the amount of noise in the interferogram and is one of the key parameters used to asses the performance of inteferometric SAR (InSAR) products. When considering a interferometric acquisition, consisting of a master acquisition $I_1$ and a slave one $I_2$, the corresponding interferometric coherence is computed as the normalized complex correlation coefficient

$$\gamma = |\gamma| \cdot e^{j\angle\gamma} = \frac{\mathbb{E}[I_1 \cdot I_2^*]}{\sqrt{\mathbb{E}[|I_1|^2] \cdot \mathbb{E}[|I_2|^2]}} \tag{61}$$

where $\mathbb{E}[\cdot]$ represents the expectation operator. The resulting feature is defined in the range from 0 (no correlation) to 1 (perfect correlation). The coherence degradation due to quantization noise $\gamma_q$, describes the specific case of quantization noise affecting a SAR interferogram and can be computed in function of the SQNR as

$$\gamma_q = \frac{SQNR}{SQNR + 1}. \tag{62}$$

## 2.4 Deep Learning-Based Denoising

In this section the fundamentals of machine learning and deep learning are briefly recalled. In particular, the main features of convolutional neural networks (CNN) are provided, together with their training. Afterwards, the concept of deep residual learning is introduced, concentrating on its particular application to denoising.

### 2.4.1 Machine Learning Fundamentals

Machine learning (ML) algorithms are designed to solve specific tasks, which typically can be categorized to be either classification or regression tasks, the latter being a generalization of the former. ML architectures are unique in that they improve their performance through experience. Usually this is implemented in the form of a learning process consisting of sample data being delivered to the algorithm, for it to process it and gain insight on an underlying mathematical model. A ML algorithm gains its experience from a so-called training-set, an ad hoc dataset used to "show" examples of the problem to be solved. The way these samples are used during the training phase also determines the type of architecture. If the training samples are shown with some context, typically in the form the expected output relative to a specific input, the training process is said to be supervised and the training data to be labelled. On the other hand, if only the inputs are shown with no information about the desired outcome, the algorithm is said to be unsupervised and the training-set to be unlabelled. A combination of both techniques also exists in the form of semi-supervised training, where an initial unsupervised training part is followed by a supervised one, used to fine-tune the performance of the system without requiring large amounts of data.

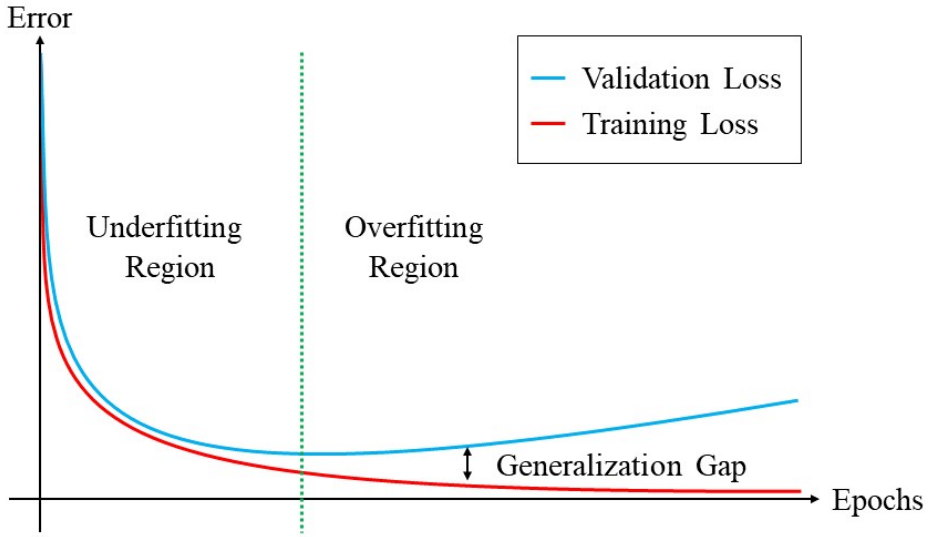



Figure 10: Typical training evolution scenario.

The performance of a machine learning algorithm is typically expressed in terms of both overall training error and generalization error, the latter consisting in the capability of the trained algorithm to possess high prediction accuracy also on previously unseen data. It is crucial for a ML system to generalize the problem presented through the training-set to new inputs. In this sense, training such an algorithm can be seen as a balancing act between diametrically opposite modelling errors, namely underfitting and overfitting. Underfitting occurs when the model generated by the ML algorithm cannot suitably capture the underlying data interrelation, missing some of the parameters required to correctly represent the problem at hand. On the contrary, overfitting consists in the opposite behaviour, with the algorithm excessively specializing on the presented training data, loosing the capability to extend the model to additional examples. This coincides with modelling some veiled noise contribution in the data in addition to the core problem itself. In accordance with the "no free lunch" theorem, no machine learning algorithm is universally better than all others, as averaged over all possible datasets, every algorithm possesses the same prediction error rate. This implies that every ML model should be designed for the specific task at hand, tuning its constituent hyperparameters through careful analyses. It is routine to keep a portion of training-set aside to optimize the hyperparameters of the model and to detect signs of overfitting in the training phase. Finally, the performance of a trained model is usually assessed using an independent, unbiased and specifically created test-set, consisting of previously unused data samples.

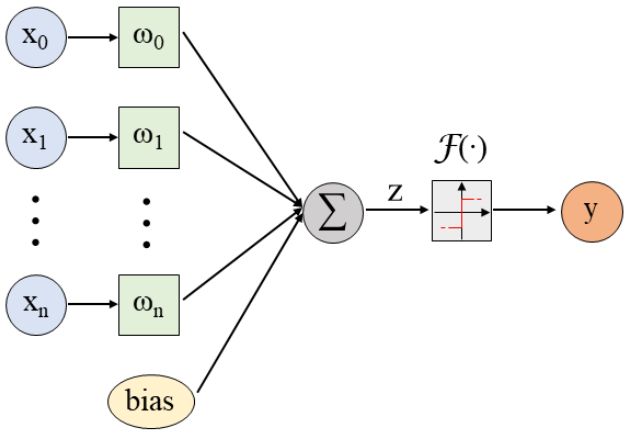



Figure 11: Structure of a perceptron.

One of the first of such a machine learning algorithms to be invented was the Perceptron. Introduced by Rosenblatt in the 1950's [19], the Perceptron is a binary classifier modelled after the biological neural networks found in animal brains. Its structure consists in a biased, linear combination $z$ of $n$ weighted inputs $x_i$, stimulating a so-called activation function $f(\cdot)$ to produce one of two possible binary outcomes:

$$y = f(z) = \begin{cases} 0 & \text{, if } z < 0 \\ 1 & \text{, if } z \geq 0 \end{cases} \text{ , } with\ z = \sum_{i=0}^{n} \{w_i x_i\} + bias \tag{63}$$

By itself, a single Perceptron is not able to solve non-linear problems such as, for example, learning a XOR function [15]. This limitation has been overcome with the introduction of the multi-layer Perceptron (MLP), which, as the name implies, consists in a multi-layered network of subsequent Perceptron.

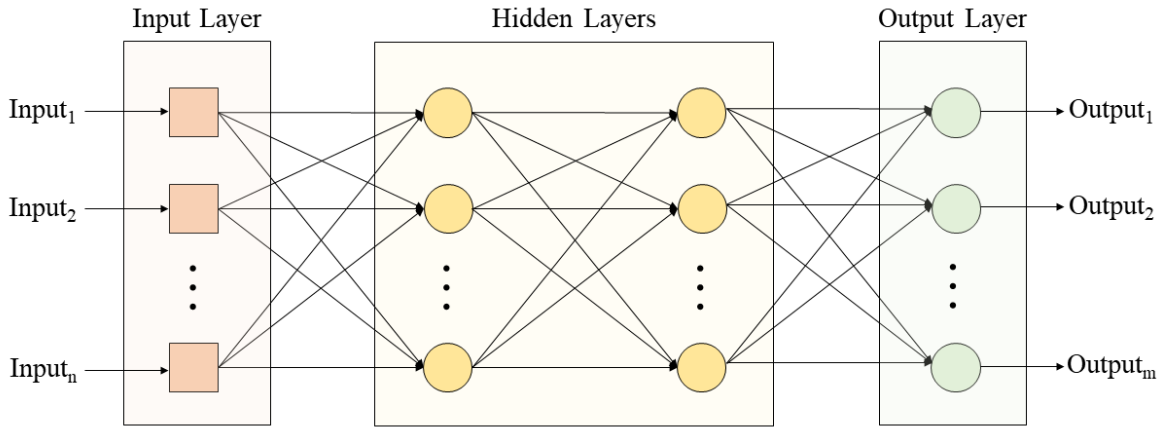



Figure 12: Structure of a MLP architecture.

The MLP is a feedforward ANN, meaning that information flows naturally only in one direction, from the input, through its inner layers, to the output. The depth of such a network is defined by the number of hidden layers interposed between the input- and output-layer. A network with many hidden layers is appropriately referred to as deep neural network. The complexity of the network combined with the usage of non-linear activation functions, is what gives these networks the capability of solving non-linear problems. One of today's most used non-linear activation functions is the Rectified Linear Unit (ReLU) [5], as it possesses several advantages over other activation functions:

- sparsity, meaning not all neurons are activated at the same time since it is capable of outputting zeros; this allows for compact input representation, simplifying the model and accelerating the learning process.

- better gradient propagation, with limited vanishing gradient problems compared to saturating sigmoidal functions.

- computational efficiency, requiring only comparison, addition and multiplication to be implemented.

The MLP constitutes the basic structure for more complex feedforward neural networks, such as convolution neural networks.

### 2.4.2 Convolutional Neural Networks

A Convolutional neural network (CNN) is a biologically inspired departure from the MLP, in that it is designed to emulate the behaviour of the visual cortex, mitigating some of shortcomings of the MLP architecture by taking advantage of spatial correlation present in natural images. CNNs are designed to exploit this spatial information by relying on matrix convolutions instead of multiplications, thus they are defined by the kernel- or filter-functions they use.

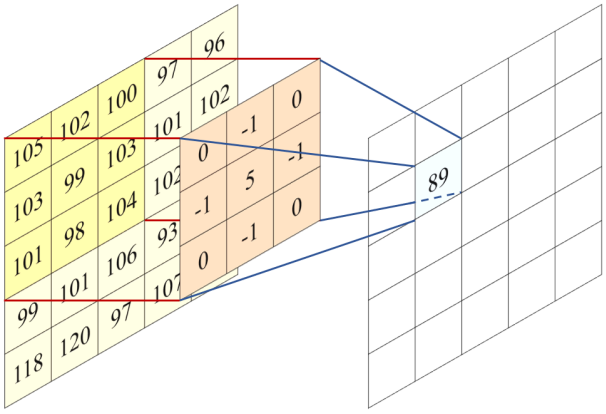



Figure 13: Representation of a 2D matrix convolution.

Similarly to the MLP, CNNs feedforward neural networks concatenate an input layer, an output layer, and a stack of hidden layers in between. These typically consist of convolutional layers, possibly followed-up by activation, pooling, normalization, and regularization layers. When compared to MLP, the following characterization can be made:

- Local connectivity, meaning that each individual neuron is connected to only a small region of the input, called the receptive field, thus responding only to localized features. The Stacking of such multiple layers allows the network to create increasingly larger representations of the input, leading to non-linear responses;

- Three dimensional neuron arrangement, with each individual column made up of neurons all connected to the same limited region;

- Shared weights, as in all neurons in the same layer share the same weight vector and bias. This results in translational equivariance, meaning that each filter is tuned to find a specific feature regardless of its position in the input.

As previously mentioned, pooling layers are often used in CNNs. These layers implement a down-sampling process, typically required to enhance the robustest to changes in the position in an image of a specific feature map, a property known as translational invariance. Effectively, the output of the

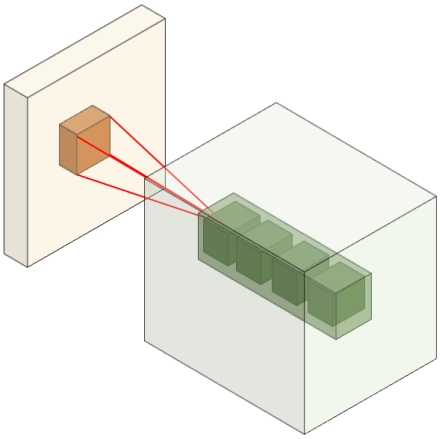


Figure 14: Representation of a 2D CNN convolution.

previous layer is substituted with a local computed statistic (e.g. the mean or the maximum), thus reducing the spatial size of the feature map and, consequently, the amount of required parameters.

Batch Normalization (BN) is another key element of deep learning architecture, which is meant to mitigate the problem of internal covariance shift which can occur during when training using mini-batches [8]. While still debated, it is believed that during training the parameter changes of preceding layers causes the means and variances of the inputs to change, thus requiring the current layer to repeatedly readjust to new distributions. The problem is compounded in deep learning architectures, with each layer amplifying the drift of the preceding one. Batch Normalization alleviates this effect by fixing the means and variances of the input to each layer, forcing them into a unit Gaussian distribution with zero-mean and unitary standard deviation. Given a mini-batch $\mathcal{B}$ of samples $\{x_1, .., x_n\}$, the resulting BN algorithm consists of the following steps:

1. The mini-batch's mean is computed as:

$$\mu_{\mathcal{B}} = \frac{1}{n} \sum_{i=1}^{m} x_i. \tag{64}$$

2. The mini-batch's variance is computed as:

$$\sigma_{\mathcal{B}}^2 = \frac{1}{n} \sum_{i=1}^{m} \left(x_i - \mu_{\mathcal{B}}\right)^2. \tag{65}$$

3. Finally,The mini-batch is normalized as:

$$\hat{x}_i = \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \varepsilon}} \tag{66}$$

   where $\varepsilon$ is used for numerical stability.

4. Finally, the mini-batch is scaled and shifted as:

$$y_i = \gamma \cdot \hat{x}_i + \beta. \tag{67}$$

where $\gamma$ and $\beta$ are the scale and shift coefficients, respectively, and are learned during the training phase. For non-Gaussian distributions, Batch Normalization is preferably applied before the activation function [8], expecting the non-linearity to produce a stable distribution. Consequently, it is possible

to substitute the linear combination defined in (63) with

$$y = f\left(\sum_{i=0}^{m}\{w_i x_i\}\right) \tag{68}$$

as the normalization process effectively eliminates the bias term.

### 2.4.3 Artificial Neural Network Training

In feedforward neural networks, such as MLP and CNN, training typically happens by firstly propagating information from input to output (forward propagation), and secondly from output to input (back propagation); the resulting algorithm takes the name of backpropagation (BP).

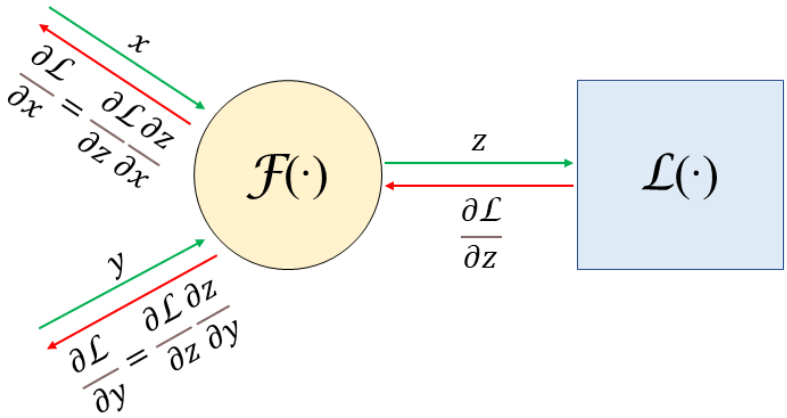



Figure 15: Example of error backpropagation for a perceptron with two inputs.

The goal of this operation is to find the global minimum in the solution space defined by the parameters of the network and the utilized loss function, representing the cost associated to the difference between the network's effective and expected outputs. Such a mapping function is defined according to the machine learning task to be solved. The amount by which each weight is updated, is determined by computing and back-propagating, one layer at a time, the gradient of the loss function with respect to each weight using the chain rule

$$\frac{\partial L}{\partial z} = \frac{\partial z}{\partial x}\frac{\partial L}{\partial z} \tag{69}$$

where $L$ is the loss function evaluated at the output $z$, and $x$ the input of the current layer. In a CNN, the chain rule is repeatedly applied to compute the gradients of the weight and input matrices in the convolutional layers. All weight coefficients in a kernel contribute to each value in the output feature map, therefore affecting the final outcome of the loss function. In order to minimize the error of the network, each weight $w_{ij}$ is updated by an amount $\Delta w_{ij}$ which is equal to the opposite of the respective gradient

$$\Delta w_{ij} = -\eta \frac{\partial L}{\partial w_{ij}} \tag{70}$$

where $\eta$ is the learning rate, which effectively represents the speed at which the optimization algorithm moves towards the local minimum. The training process loops, continuously updating the network's parameters until either the global minimum has been reached or a predetermined stopping condition has been satisfied. There are different gradient-based optimization algorithms which can be used during backpropagation. A commonly used example is Stochastic gradient descent (SGD), which replaces the actual computation of the gradient with an estimation, achieving faster iterations at the cost of slower convergence-rate. A critical hyperparameter to be tuned in SGD is the learning rate, as this is fixed and therefore has a great impact on the resulting network performance. Optimizers such as Adam (Adaptive Moment Estimation) overcome this limitation by dynamically updating the

learning rate of the model.

### 2.4.4 Deep Residual Learning

A residual neural network (ResNet) is a family of artificial neural networks that diverge from non-residual ones due to the presence of skip connections (or shortcuts) to bypass some of the layers [7]. These skip connections provide a solution to two important problems in deep neural networks:

- the problem of vanishing gradients
- the degradation of the training error

It has been observed that for "plain" neural networks, deeper architectures tend to suffer from performance degradation when compared to shallower ones [7]. The introduction of shortcuts shortens the path from input to output, effectively simplifying the network, making it less vulnerable to perturbations which would require excessive amounts of training samples to cancel out.

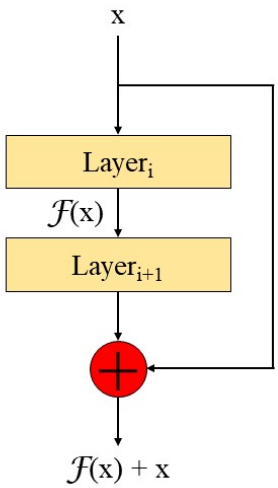



Figure 16: Basic residual learning block.

Different residual block configuration exist, each one more or less suitable for different kinds of applications. Considering a stack of convolutional layers, a skip connection is introduced from the input of the stack to its output. If the input and output dimension of the residual block are the same, the bypass is said to be an identity shortcut connection:

$$y = \mathcal{F}(x, \{\mathcal{W}_i\}) + x \tag{71}$$

Such an identity shortcut does not introduce any further parameter to the network, since the information of the bypassed data is preserved across different layers. In case of a mismatch, a linear projection $\mathcal{W}_s$ must be performed to match the two dimensions. The resulting bypass connection is thus called projection shortcut:

$$y = \mathcal{F}(x, \{\mathcal{W}_i\}) + \mathcal{W}_s x. \tag{72}$$

The resulting network architecture has been shown to effectively improve both convergence speed and performance, making it an attractive proposition for very deep architectures.

### 2.4.5 Deep Neural Networks for Image Denoising

Image denoising aims at recovering a clean image $\boldsymbol{x}$ from a noisy measurement $\boldsymbol{y}$, which is distorted by a degradation term $\boldsymbol{\eta}$ according to

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\eta}. \tag{73}$$

In a statistical framework, a successful denoising operation can be implemented when accurate statistical descriptions of the signal and noise components are available. A very common assumption

is that the noise term $\boldsymbol{\eta}$ is an additive white Gaussian noise (AWGN) process. Conversely, various models have been proposed to model natural images, including Gaussian scale mixtures (GSM) based on multi-scale wavelet decomposition [18], and Markov random field (MRF) models [21].

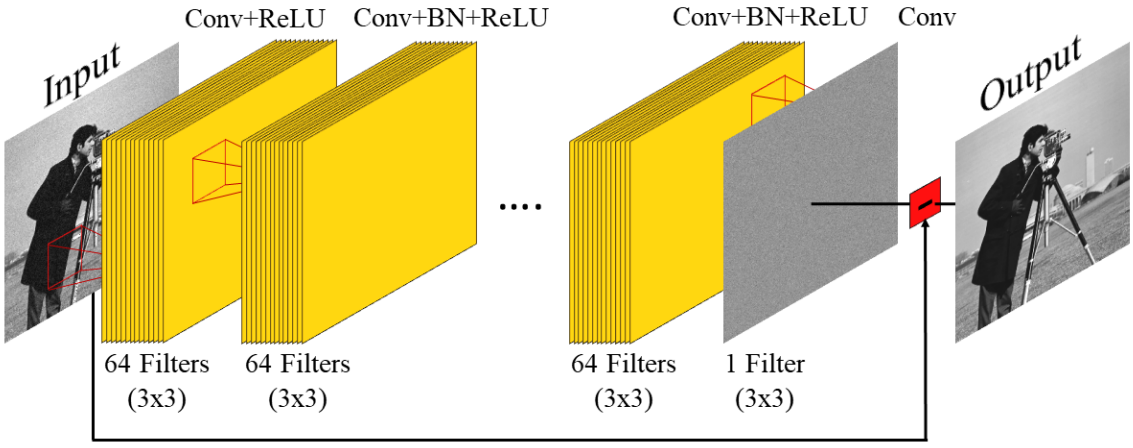



Figure 17: Structure of the DnCNN architecture proposed in [22].

These high-performance denoising algorithms typically suffer from excessive complexity, as they require extensive, time consuming computations for the optimization of their hyperparameters. This in turn limits the achieved performance levels when compared to the theoretical indications [9]. The adoption of convolutional neural networks essentially changes the underlying approach towards denoising, presenting it in terms of regression rather than a statistical modelling problem. Furthermore, very deep architectures posses great representational power and can be effectively trained exploiting the latest advancements made in field. The latter primarily consisting of the ReLU activation function discussed in section 2.4.1, the batch normalization algorithm discussed in section 2.4.2, the Adam optimizer introduced in section 2.4.3, and the deep residual learning architecture covered in section 2.4.4. In [22] these improvements have been jointly extended for the first time to the problem of denoising of natural images. The proposed DnCNN architecture treats the denoising process as a discriminative learning problem, effectively teaching the network how to discern the noisy component $\boldsymbol{\eta}$ from the underlying clean image $\boldsymbol{x}$. In other words, the architecture takes advantage of the deep residual learning scheme to predict the residual image $\hat{\boldsymbol{\eta}}$ and subtracting it the bypassed noisy observation $\boldsymbol{y}$, obtaining

$$\hat{\boldsymbol{x}} = \boldsymbol{y} - \hat{\boldsymbol{\eta}}. \tag{74}$$

As it will be presented in chapter 4, the residual learning paradigm is the basis on which we developed the proposed approach for the removal of quantization noise from focused SAR data.

# 3 Analysis of Quantization Errors in Simulated and Real SAR Data

Starting from an overview of the state-of-the-art quantization methods used by today's SAR systems, benchmarks of the most used techniques are presented and evaluated in this chapter. Finally, a novel approach to quantization resource optimization is introduced, exploring the feasibility of quantization noise predictability for scene-specific bit-rate allocation.

## 3.1 State-of-the-Art Quantization Methods for SAR

During a data-take, SAR systems need to store the received echoes from the target-scene on board. This requires both the analog-to-digital conversion of the backscattered signal and the compression of the resulting digital representation to meet the stringent on-board memory limits.

In this context, the Block-Adaptive Quantization (BAQ) scheme is one of most widely recognized methods for raw data compression, since it represents a good compromise between data compression-rate and implementation complexity [13]. Introduced in Section 2.2.4, BAQ is a quantization technique that reduces the number of required quantization bits by adapting the dynamic range of the quantizer according to the local statistics of the signal. As the local dynamic range of is always equal or smaller than that of the entire signal [20], this approach results performance improvement. Over the years much effort has gone into researching novel quantization strategies, aiming at fine-tuned resource and performance optimization. FFT-BAQ is natural evolution of the original algorithm, as it consists of the BAQ scheme operating in the frequency domain. It works by exploiting the spectral diversity introduced in the elevation and azimuth dimensions of the SAR acquisition. It achieves approximately 5 dB of improvement over the time domain implementation, at the expense of significantly higher computational requirements [4].

In Block-Adaptive Vector Quantization (BAVQ) a vector quantization step follows the BAQ compression scheme, allowing to further increase the data-rate compression by encoding a vector of BAQ-compressed samples using codewords from a pre-defined, scene independent codebook. While performance gains of up to 3 dB have been achieved, the higher complexity of the compression scheme has hindered its diffusion [16]. The Flexible Dynamic BAQ (FDBAQ) represents a relatively recent

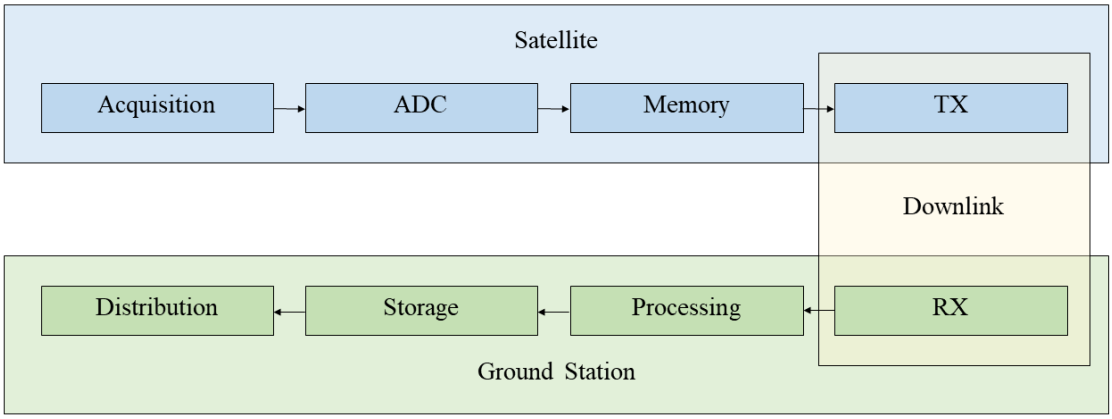



Figure 18: Block diagram of a typical spacerborne remote sensing pipeline, from acquisition to distribution.

method aimed at improving quantization performance and is used on-board ESA's Sentinel-1 satellites. It works by adaptively selecting the bit-per-sample rate of the quantizer according to the locally estimated signal-to-thermal-noise ratio (STNR), resulting in an overall average SQNR improvement of 0.5 dB compared to a fixed-rate BAQ, with peaks of 4 dB for very bright targets [2]. Such an adaptive technique represents a first step toward a target-dependent rate allocation. While compression techniques like FDBAQ take advantage of statistical knowledge of the acquired raw data to optimize the data-rate, on-board architectures inevitably lack knowledge about the impact the quantization process has on the resulting focused SAR image.

## 3.2 Benchmarks of Quantization Methods on Simulated Data

In this section, the performance of the different quantization methods introduced in Section 2.2 are analyzed and compared for different bit-rates, considering the performance metrics introduced in Section 2.3.4.

### 3.2.1 Simulated Data

In SAR systems, the on-board quantization process typically happens under strict memory and computation limitations, especially for spaceborne systems. To test the trade-offs between some of the optimal and sub-optimal quantizers used on board, the performance of each implementation has been evaluated for synthetic SAR raw data. As a consequence of the central limit theorem, the in-phase (I) and quadrature (Q) components of the received SAR echos can be modelled as independent, stationary, zero-mean Gaussian processes [3]. Thus, a complex, Gaussian-distributed, zero mean signal of unitary variance has been considered for the following performance tests (Figure 19). According to this assumption, the performance evaluation has been limited to consider only granular, overhead, and low amplitude suppression errors, ignoring the effects of low scatterer suppression described in Section 2.3.3, as these become apparent only after the SAR focusing process.

### 3.2.2 Uniform ADC Performance

Quantization performance testing starts with the uniform quantizer, it being the simplest possible implementation. As introduced in Section 2.2.2, a uniform quantizer first clips the dynamic range of the input signal $x[n]$ to the $\pm V_{clip}$ range, then discretizes its values by assigning each DT value to one of a finite set of equally spaced representations $x_q[n]$ covering the entire dynamic range, resulting in the characteristic stair-stepped input-output relation represented in Figure 7. For high enough

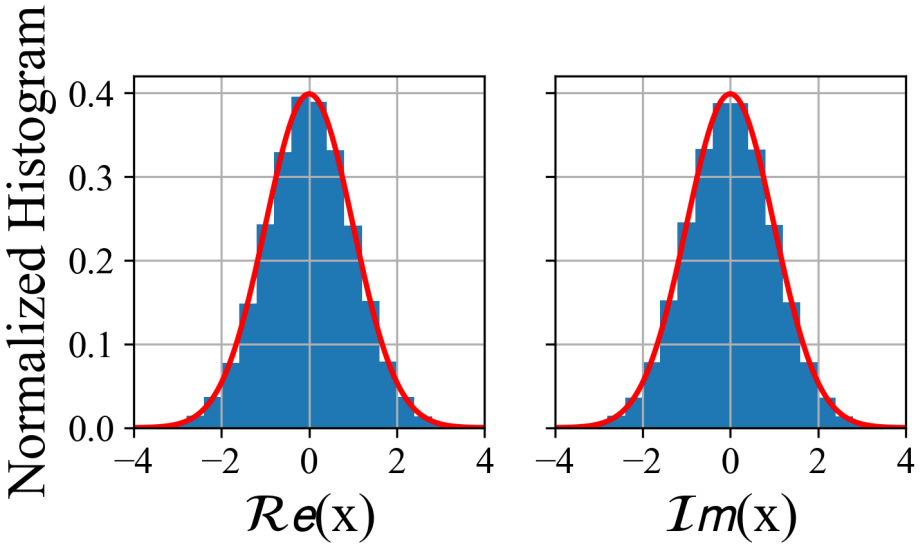



Figure 19: Distribution of the complex Gaussian signal used for performance evaluation of the different quantization techniques.

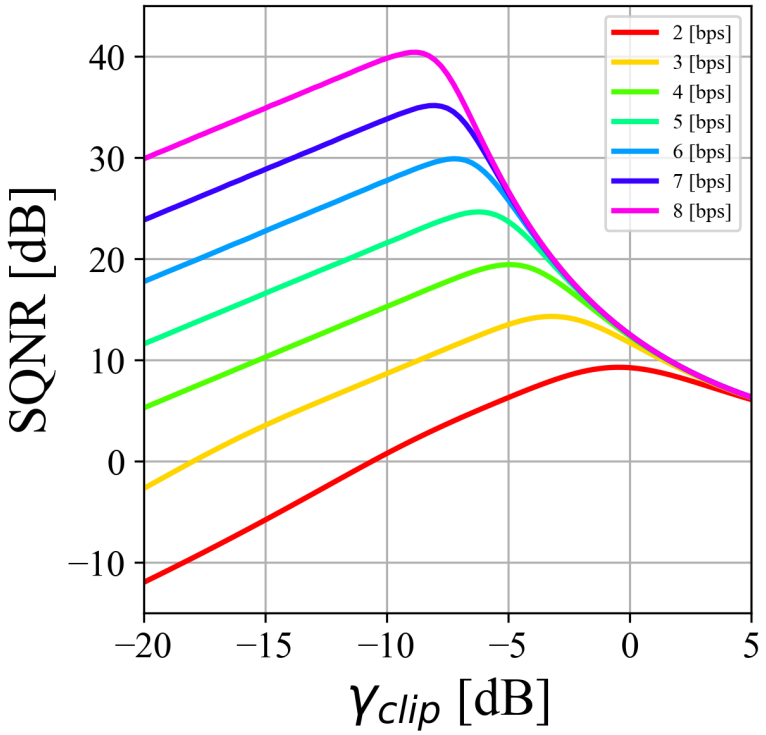



Figure 20: Signal-to-quantization-noise ratio (SQNR) of a uniformly quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bits-per-sample.

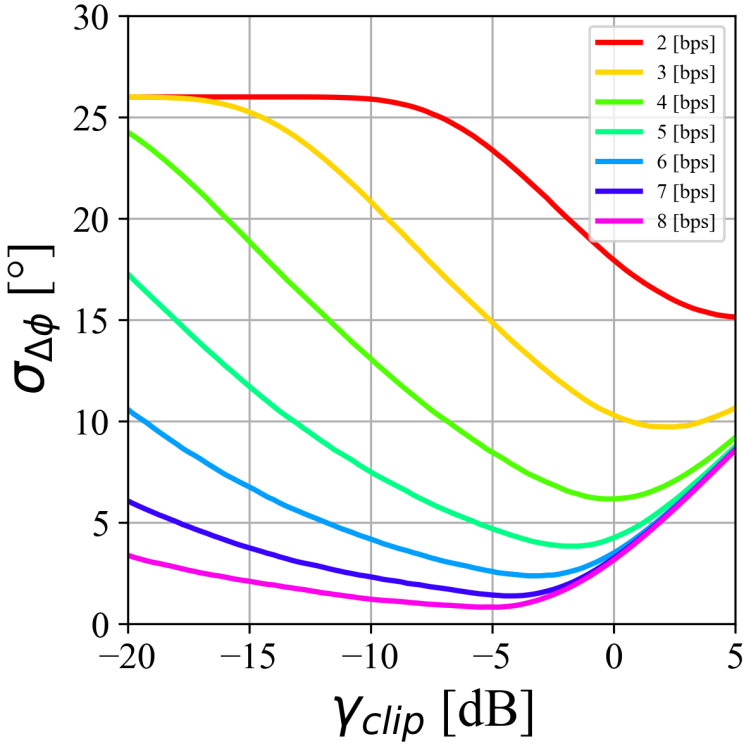



Figure 21: Phase error standard deviation ($\sigma_{\Delta\phi}$) of a uniformly quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates.

bit-rates, it is possible to assume that the granular error for such a quantizer is going to be uniformly distributed between $\pm\Delta$, where the step-size $\Delta$ is the distance between two quantization intervals. For the performance evaluations a midtread uniform quantizer has been considered, as it is used on board the TanDEM-X constellation. The first performance experiment consists in evaluating the signal-to-quantization-noise ratio (SQNR) of the uniformly quantized complex signal for different signal-to-clipping ratios $\gamma_{clip}$ and using different numbers of bits-per-sample. We recall that the signal-to-clipping ratio, as defined in Section 2.3.4, relates the standard deviation of the input signal $\sigma_x$ to the clipping value $V_{clip}$ of the quantizer, hence higher values of $\gamma_{clip}$ correspond to dominant overload errors while lower ones are consistent with increased distortions caused by granular error. Let us consider a $\gamma_{clip}$ interval from -20 dB to 5 dB, and seven possible bits-per-sample values from 2 to 8 bps. The resulting SQNR performance plot is represented in Figure 20.

Comparing the different bit-rates, one can see that the performance gain is consistent with the expected 6 dB/bit uplift described in (59), with higher bits-per-sample allowing to discretize the input signal with incrementally less distortion, resulting in a higher SQNR. Furthermore, it is possible to notice the aforementioned impact of granular and overhead errors when considering different values of $\gamma_{clip}$. Depending on the utilized number of bps, $\gamma_{clip}$ values above -8 dB to 2 dB for 8 bps to 2 bps respectively, show a clear performance deterioration due to a decreasing clipping value, while for $\gamma_{clip}$ values below -10 dB and -2 dB the performance degradation is dominated by granular error. It is possible to observe that in a typical SAR acquisition scenario, the clipping value $V_{clip}$ used by the quantizer is fixed, while the standard deviation of the acquired signal is heavily localized and dependant on the imaged target scenario. This entails that the performance of a uniform quantizer is severely limited both by it being sub-optimal for normally distributed signals and by adopting a set of fixed representation values. Another important performance metric, especially for interferometric applications, is the phase difference defined in (60). The impact in terms of phase error standard deviation $\sigma_{\Delta\phi}$ is plotted in figure 21 for different bits-per-sample and for different values of $\gamma_{clip}$. As expected, using fewer bits has a significant impact on the overall performance. The best possible performance for each bps value comes as a $\gamma_{clip}$ trade-off between the granular and clipping errors. Finally, the generic

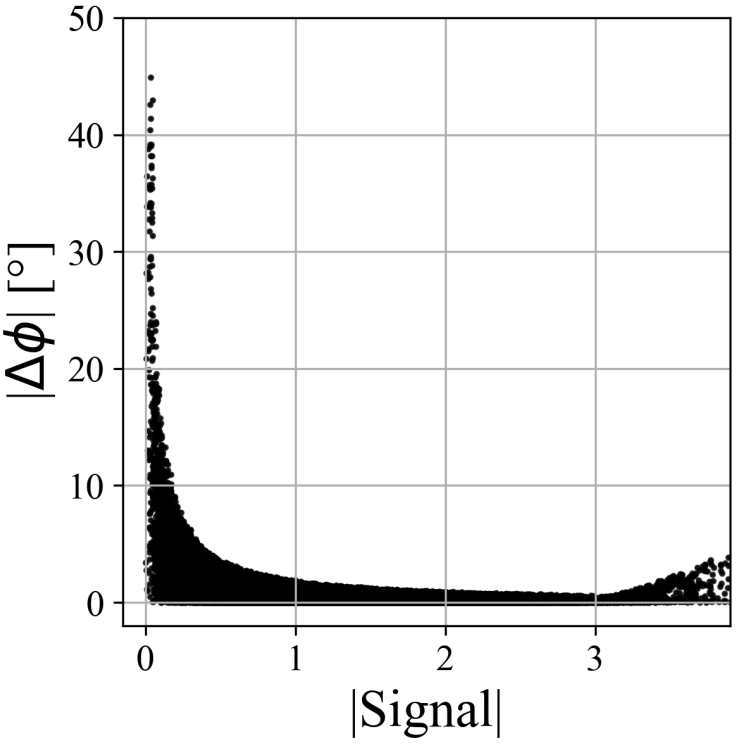



Figure 22: Phase error magnitude ($|\Delta\phi|$) of a uniformly quantized, normally distributed signal for different input signal magnitudes.

| Compression Rate | $C$ Value | $E_{max}$ Value | $M_{max}$ Value |
|---|---|---|---|
| 8:2 | 2.20374 | 24 | 1 |
| 8:3 | 5.28038 | 20 | 3 |
| 8:4 | 8.50475 | 16 | 7 |
| 8:5 | 11.90135 | 12 | 15 |
| 8:6 | 15.25490 | 8 | 31 |
| 8:7 | 18.44601 | 4 | 63 |

Table 3: BAQ parameter table extended by the missing 8:5 bit and 8:7 bit entries in the TanDEM-X mission table 2.

distortion effects of low amplitude suppression introduced by Cartesian quantizers are explored. As described in Section 2.3.2, the phase error magnitude $|\Delta\phi|$ resulting from the quantization process is inversely proportional to the magnitude of the input signal, with the overall error tending to 45° as the signal magnitude tends to zero. Considering a 8-bit uniform quantizer with a fixed clipping value of $V_{clip} = 3$, the resulting relationship is shown in Figure 22, where each dot represents the amplitude error resulting from a single signal sample. For an input signal magnitude comprised in the $[0; 3]$ range, it is clear to observe both the expected inverse relationship between phase error magnitude and signal magnitude, and the upper bound of 45° for the phase error magnitude. Interestingly, it is also possible to observe the impact of the overhead error, as for signal magnitude values larger than $V_{clip} = 3$, the introduced distortion increasingly affects the phase error.

### 3.2.3 BAQ Performance

Unlike the uniform quantizer, the Block-Adaptive Quantization (BAQ) algorithm uses and adaptive quantization scheme. As such, it and its evolutions have become the compression algorithm of choice for many SAR systems such as the TanDEM-X [10], Sentinel-1 [2], and COSMO-SkyMed constellations [1].

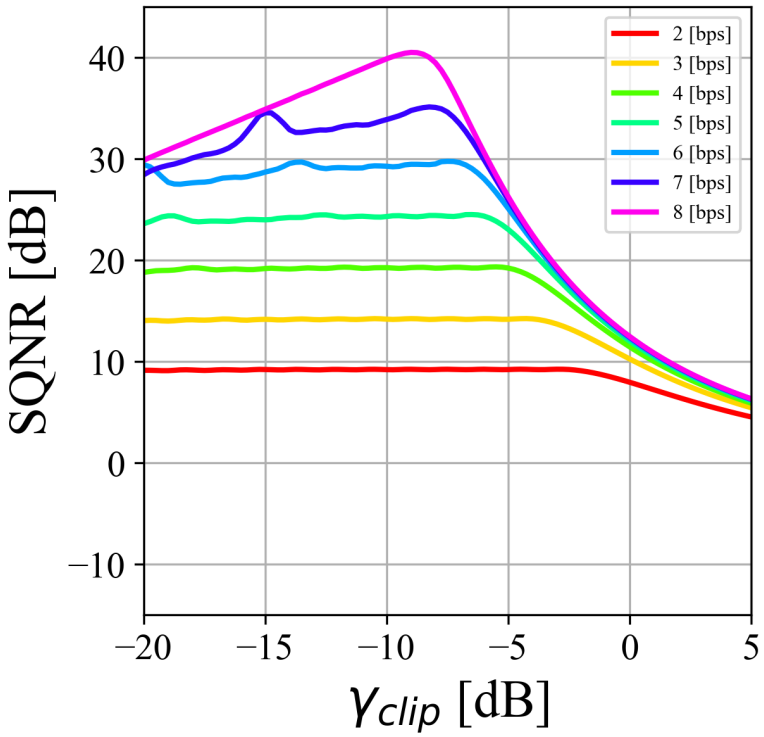



Figure 23: Signal-to-quantization-noise ratio (SQNR) of a BAQ quantized, normally distributed signal at different signal-to-clipping ratios $\gamma_{clip}$ and for different bit-rates.

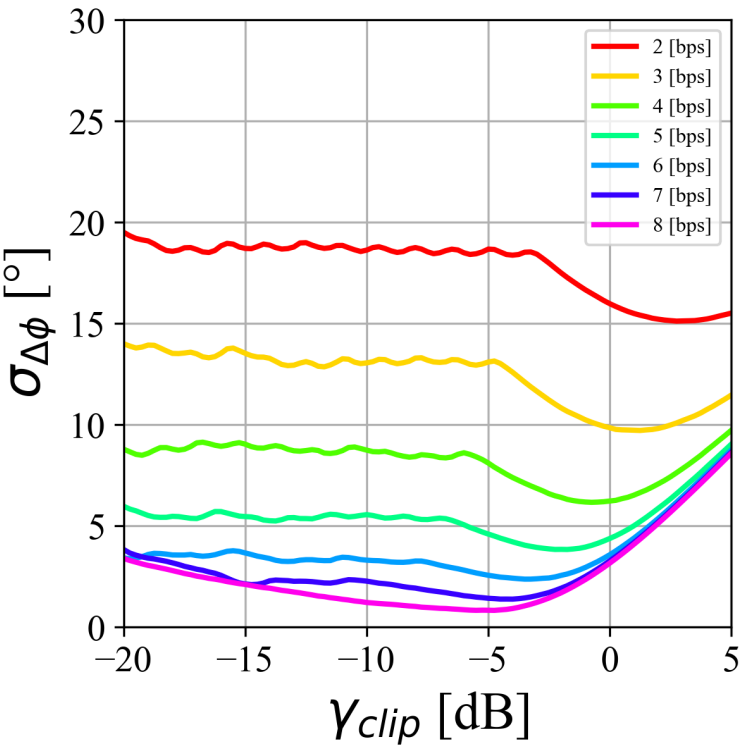



Figure 24: Phase error standard deviation ($\sigma_{\Delta\phi}$) of a BAQ quantized, normally distributed signal at different signal-to-clipping ratios $\gamma_{clip}$ and for different bit-rates.

In order to extend the experiments defined in Section 3.2.2 to the BAQ compression scheme, firstly the TanDEM-X quantization parameters defined in Table 2 have been extended to cover the tested [2; 8] bps range, resulting in the compression coefficients described in Table 3. Note that a 8:8 compression rate implies a bypass of the BAQ compression scheme, resulting in a uniform 8-bit quantization. Firstly, the BAQ performance degradation is tested for different bit-rates and for different signal-to-clipping ratios (Figure 23). Contrary to was has been seen for the uniform quantizer, the BAQ scheme displays virtually constant levels of distortion when granular noise constitutes the dominant error contribution. This confirms the expectation that the algorithm can adapt the set of available sample reconstructions $x_q[n]$ to the statistics of the quantized block. When the overload error is the dominant distortion component a significant portion of distribution tails is clipped, thus both BAQ and uniform ADC have comparable performance. Finally, it can be seen that the performance achieved by the 8-bit uniform ADC represents the upper bound for the achievable performance as in the TerraSAR-X and TanDEX-X systems, the BAQ algorithms implements a more elaborated compression step after the 8-bit uniform quantizer. In terms of phase error standard deviation, Figure 24 confirms that by adapting the sample location to the dynamic range of the input signal, the BAQ algorithm shows near constants performance for values of $\gamma_{clip}$ in which granular noise is dominant.

| $\mathbf{2}^{N_b}$**-samples** | $\mathbf{x}_q$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 4 | 0.4528 | 1.5104 | | | | | | |
| 8 | 0.2451 | 0.7560 | 1.3439 | 2.1519 | | | | |
| 16 | 0.1284 | 0.3880 | 0.6568 | 0.9423 | 1.2562 | 1.6180 | 2.0690 | 2.7326 |
| 32 | 0.0659 | 0.1980 | 0.3314 | 0.4667 | 0.6049 | 0.7471 | 0.8946 | 1.04879 |
| | 1.2118 | 1.3863 | 1.5762 | 1.7872 | 2.0287 | 2.3177 | 2.6911 | 3.2607 |

Table 4: Max-Lloyd reconstruction values as obtained by the iterative algorithm.

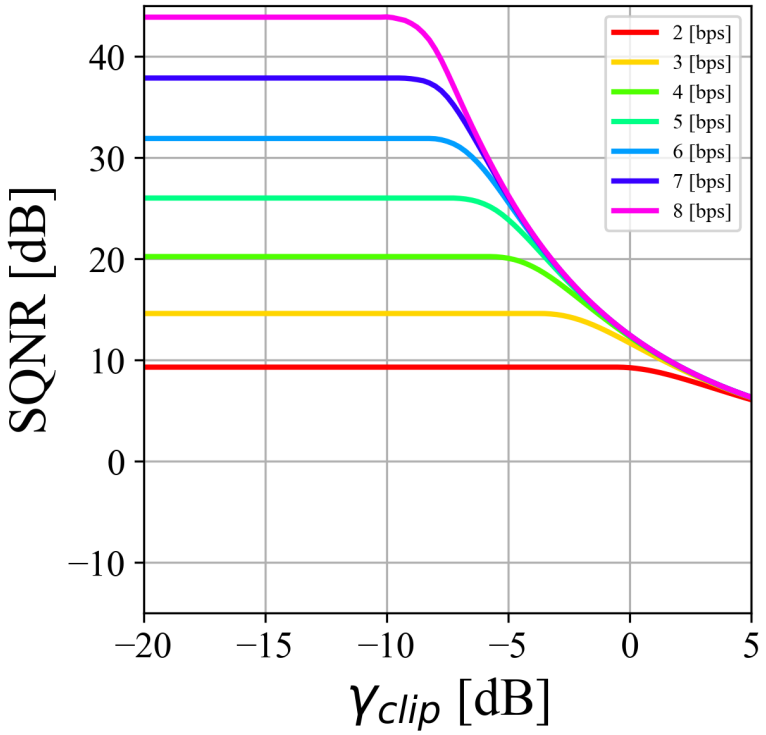



Figure 25: Signal-to-quantization-noise ratio (SQNR) of a Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates.

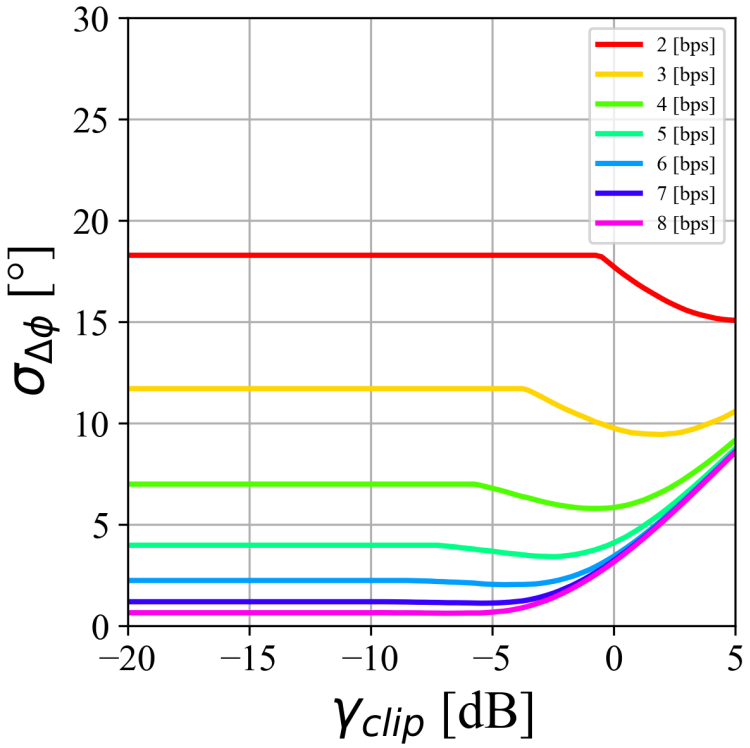



Figure 26: Phase error standard deviation ($\sigma_{\Delta\phi}$) of a Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates.

### 3.2.4 Max-Lloyd Quantizer Performance

The Max-Lloyd algorithm defined in Section 2.2.3 is a non-uniform, optimum quantizer since it finds the optimum decision and reconstruction levels that minimize the overall mean squared error (MSE) [14]. The algorithm has been implemented by means of an iterative process. Starting from an initial random guess, the process alternatively predicts the optimum decision levels (38) and reconstruction levels (37). The algorithm has been applied considering the Gaussian signal described in Section 3.2.1, for the tested [2; 8] bps range. The resulting reconstruction values are displayed for $N_b = 2, 3, 4, 5$ bits in Table 4. The good agreement of these results with those numerically found in [14] proves the level of precision achievable by the iterative process. Figure 25 shows the performance of the optimum quantization algorithm for different values of $\gamma_{clip}$ and different bit-rates. It is possible to notice that for $\gamma_{clip}$ values below -10 dB and -2 dB, at 8 bps and 2 bps respectively, there is no performance degradation due to granular error, as the algorithm optimally adapts the location of the reconstruction samples to the pdf (i.e. the expected value) of the input signal. The overload error still limits the performance for stricter clipping values $V_{clip}$ as a bigger portion of the tail of the input distribution falls outside the dynamic range supported by the quantizer. Similarly, testing the performance of the Max-Lloyd quantizer in terms of phase error standard deviation (Figure 26) shows that the algorithms overcomes the performance limit imposed by the 8-bit uniform quantization step in the on-board TanDEM-X BAQ scheme.

### 3.2.5 Quantizer Performance Comparison

To better assess the relative performances of the tested quantization algorithms, the previously shown results have been condensed in individual plots. Figure 27 shows the absolute performance values in terms of signal-to-quantization-noise ratio for different values of signal-to-clipping ratio. As expected, the results show that the Max-Lloyd quantizer represents the upper limit for the performance achievable by competing quantization techniques. The BAQ compression algorithm is indeed sub-optimal, but it achieves significant performance gains over the uniform quantizer, especially for signal distri-

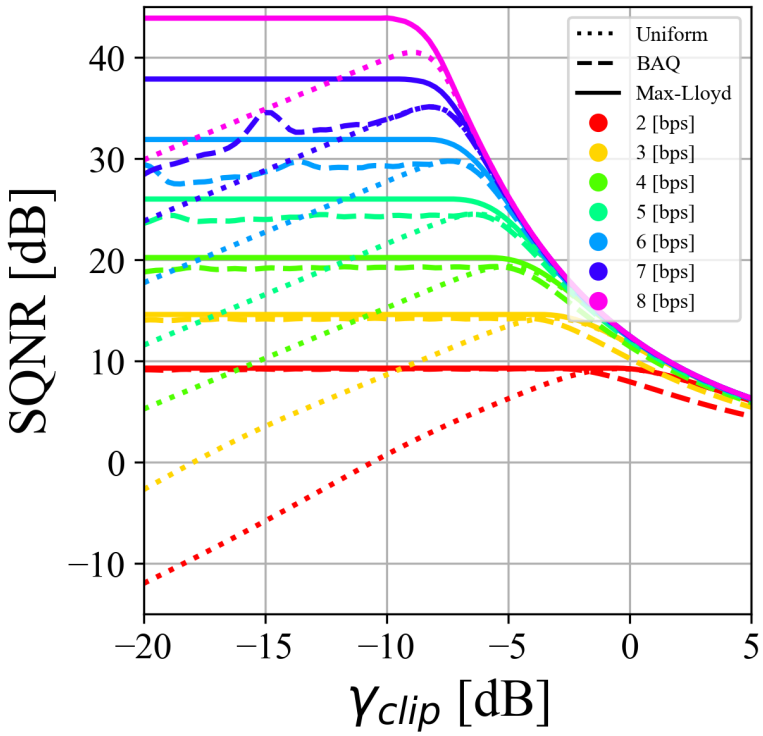



Figure 27: Signal-to-quantization-noise ratio (SQNR) comparison of a uniformly, BAQ and Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates.

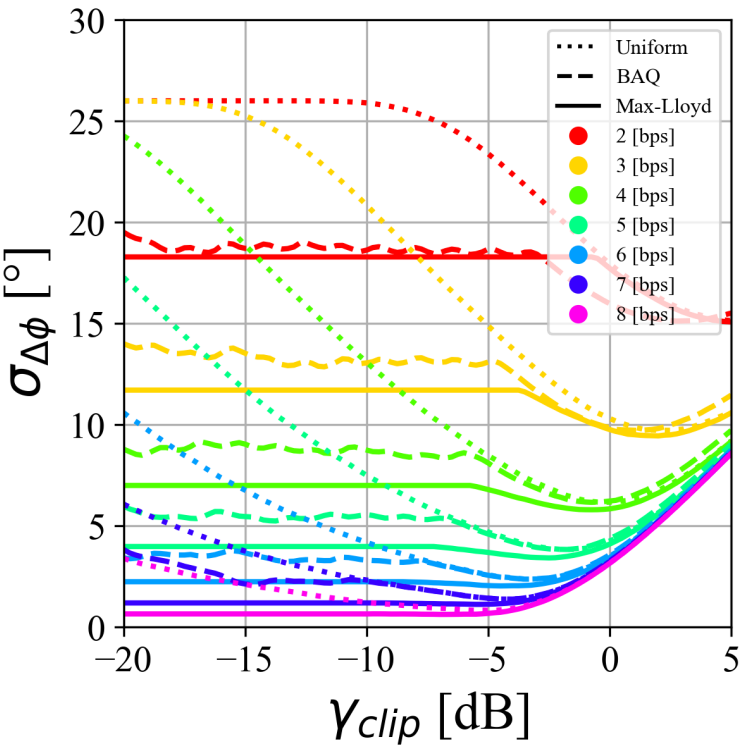



Figure 28: Phase error standard deviation ($\sigma_{\Delta\phi}$) comparison of a uniformly, BAQ and Max-Lloyd quantized, normally distributed signal at different signal-to-clipping ratios ($\gamma_{clip}$) and for different bit-rates.

butions with different dynamic ranges compared to a chosen fixed clipping value. It is also worth noting that BAQ scheme has the biggest impact for the lowest numbers of bits-per-sample, as the overall achievable performance is upward-limited by the degradation introduced by the 8-bit uniform quantizer. Similarly, testing the phase error standard deviation (Figure 28) shows a significantly lower signal degradation using the BAQ scheme over the uniform quantizer. The optimum quantizer represents the lower bound to the achievable standard deviation. As the overhead error becomes the dominant quantization noise component, the performance of all three methods converge and degrade as they become more and more limited by the dynamic range available to the quantizers. It is of significant importance to keep in mind that the optimum quantizer requires more computational resources than either the uniform and the BAQ quantization schemes. Therefore, the BAQ algorithm represents a powerful trade-off between performance and complexity, justifying its widespread use on-board memory- and computationally-constrained remote sensing platforms.

## 3.3 Preliminary Analysis on Real SAR Data for Data-Rate Optimization

The following section qualitatively tests the possibility of predicting the impact of a scene on the overall quantization error. This is analysis is contextualized as a potential tool for ad hoc, scene-specific resource allocation optimization. The evaluated performance metrics are tested for linear correlation with the quantization error present in each of the utilized real focused TanDEM-X SAR acquisitions.

### 3.3.1 Concept

In recent years, fine-tuned performance and resource optimizations have been at the center of SAR raw-data compression research. So far, improving upon fixed-rate BAQ mainly involved better exploitation of the entropy in the raw-data working in transform-domains, exploiting vector-encoding and exploring other strategies. ESA's FDBAQ represents in some sense a novel approach to SAR raw-data quantization as it dynamically allocates more or less resources to portions of the signal based on its statistics. This approach is still somewhat limited, in that the optimization process is done on the raw-data, with no information about the overall data volume and the final impact on the focused products. In this section we explore the idea of predicting the error magnitude in focused SAR imagery for a given data take, using previous acquisitions of the same area. Such a tool would allow for on-ground performance optimization by determining the most suitable resource allocation for an expected level of performance. This process in limited in its nature as it requires the optimized bit-rate allocation map to be uplinkend to the satellite for commanding. On the other hand, this potentially allows pre-existing systems, like TanDEM-X and TerraSAR-X, to take advantage of this improvement without requiring expensive or otherwise unfeasible hardware upgrades. The analysis of the predictability of error magnitude will be mostly qualitative. Starting by modelling the impact of the quantization error generated in the raw-data domain on the focused SAR data, the Pearson correlation coefficient will be used to asses linear relationships between the error magnitude and locally computed image statistics.

### 3.3.2 Used Dataset

The dataset used for the noise prediction experiments consists of six TanDEM-X, single-look, slant-range, SAR images representing a heterogeneous mix of different scenes:

- Antarctica: snow and ice coverage with mountainous topography.
- Brazil: hilly, rain-forest with clear-cuts.
- Greenland: snow- and ice-covered flatlands.
- Iowa: agricultural area.
- Mexico City: heterogeneous urban area with sloping, rocky areas and water bodies.
- Malaysia: mountainous tropical jungle.

The backscatter intensities of the acquisitions have been plotted in Figure 29. It is apparent that the dynamic ranges, the variances and the intensities of the images are highly diversified. As a consequence of the signal-dependent and non-linear nature of the quantization error, areas with larger dynamic ranges and standard deviations are more affected than more homogeneous and uniform areas. For all acquisitions, a high quality 8-bit and a 2-bit BAQ-compressed SAR images are available to test the extent of the quantization error in the focused-data domain. All images are Single Look Complex (SLC) products. The physical acquisition parameters and imagery statistics are shown in table 5, which lists the acquisition dimensions, the pixel-spacing (i.e. the number of meters corresponding to one pixel in the raster) and the incidence angles for the different acquisitions.

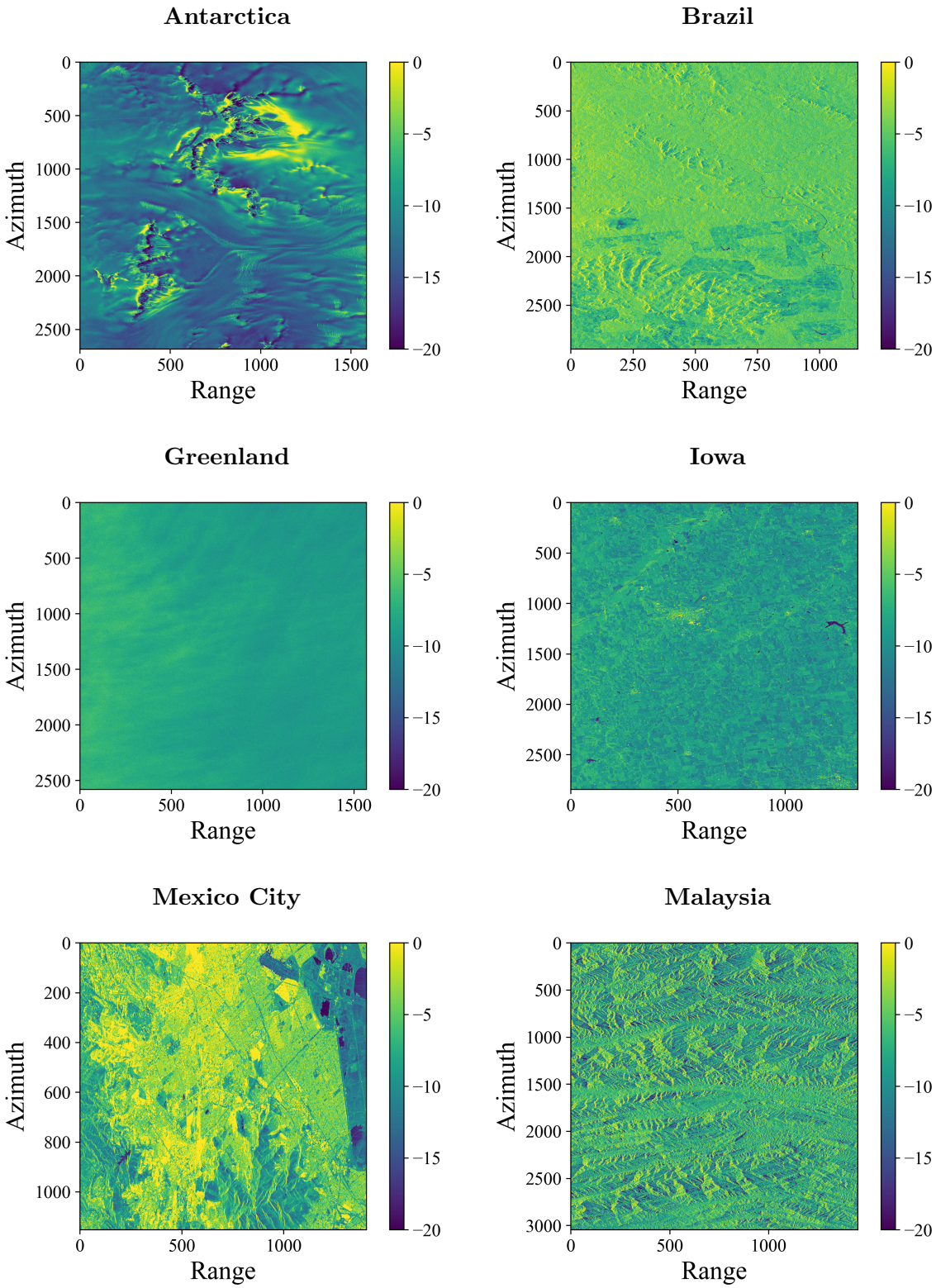



Figure 29: Intensity images of the utilized TanDEM-X acquisitions, represented in dB scale.

### 3.3.3 Experiments and Results

The core principle behind the proposed quantization error prediction, is the assessment of the impact that individual raw-data samples have on the focused image. From what has been discussed in Section 2.1.2, it emerges that each pixel in the focused image domain is dependent on all raw data samples that fall within an area of range- and azimuth-time extent corresponding to that identified by the range

| | Incidence Angle [°] | Azimuth Extension [px] | Range Extension [px] | Azimuth Pixel Spacing [m] | Range Pixel Spacing [m] |
|---|---|---|---|---|---|
| **Antarctica** | 36.2 | 26854 | 15870 | 2.2641039 | 1.3641052 |
| **Brazil** | 29.8 | 29508 | 11542 | 2.0604894 | 1.3641052 |
| **Greenland** | 40.6 | 25814 | 15692 | 2.0250783 | 1.3641052 |
| **Iowa** | 33.8 | 28438 | 13388 | 2.1379514 | 1.3641052 |
| **Mexico City** | 33.6 | 11534 | 14070 | 1.9719614 | 1.3641052 |
| **Malaysia** | 47.6 | 30456 | 14516 | 1.9963067 | 1.3641052 |

Table 5: TanDEM-X dataset acquisition parameters.

| | Azimuth Window | | Range Window | |
|---|---|---|---|---|
| | **m** | **px** | **m** | **px** |
| **Antarctica** | 4102.472 | 1812 | 8039.726 | 5894 |
| **Brazil** | 3815.008 | 1852 | 7315.976 | 5363 |
| **Greenland** | 4360.144 | 2153 | 7245.253 | 5311 |
| **Iowa** | 3983.867 | 1863 | 7591.8181 | 5565 |
| **Mexico City** | 3974.604 | 2016 | 7000.862 | 5132 |
| **Malaysia** | 4909.566 | 2459 | 7087.292 | 5196 |

Table 6: SAR focusing correlation window for the different TanDEM-X acquisitions.

and azimuth resolutions of an equivalent SLAR system, transmitting a windowed, sinusoidal impulse. Since the raw-data acquisition can be seen as an inverse focusing process starting from the focused image to produce the raw-data matrix, one can conclude that each raw-data sample is the coherent superposition of the backscattered signals from a region in the image domain coinciding with half the pulse length and the antenna footprint in the range and azimuth directions, respectively. Thus, the local raw-data statistics which determine the outcome of the BAQ process, depend on a well defined portion of the acquired scene. Finally, due to the focusing process, the resulting quantization error is coherently distributed in an area of the focused image which is proportional to half the pulse length in the range direction and the antenna footprint in the azimuth. Contextualizing this information for the available TanDEM-X dataset, the acquisition parameters reported in tables 1 and 5 can be used to determine the extent of the focusing window affecting the quantization error magnitude and distribution in the focused image. Since for all acquisitions, the the x- and y-axes correspond to the range and azimuth directions respectively, both naming schemes will be used interchangeably depending on the considered domain. The focusing-related window dimension along the x-axis $L_{sr}$ is computed as

$$L_{sr} = \frac{c_0 \tau_p}{2} = \frac{c_0}{2} \frac{\delta_{\%}}{PRF}, \tag{75}$$

where $\delta_{\%}$ is the transmission duty cycle. Note that $L_{sr}$ coincides with the distance covered by half the transmitted windowed waveform, to account for the dual travel path of wave from the transmitter, to the target and back to the receiver. The focusing-related window dimension along the azimuth direction $L_{az}$ is computed as

$$L_{az} = \lambda \frac{R_0}{L_a} = \frac{c_0}{f} \frac{h_{sat}}{\cos \theta_{inc}} L_a^{-1}. \tag{76}$$

The corresponding window dimensions in the image domain are obtained by dividing the results from (75) and (76) by the corresponding pixel-spacings as follows

$$L_x \approx \frac{L_{sr}}{k_{sr}} \tag{77}$$

$$L_y \approx \frac{L_{az}}{k_{az}} \tag{78}$$

where $k_{sr}$ and $k_{az}$ are the range and azimuth pixel-spacing coefficients, respectively. The resulting window extensions are reported in Table 6, both in meters and in equivalent number of pixels using

the previously reported range and azimuth pixel-spacings. This information is used to asses the degree

| | Average | Standard Deviation | Skewness | Kurtosis |
|---|---|---|---|---|
| | **Error Magnitude** | | | |
| **Antarctica** | 0.45108 | 0.58670 | 0.00241 | 0.00007 |
| **Brazil** | -0.04316 | 0.03235 | -0.00578 | -0.00084 |
| **Greenland** | -0.03943 | 0.14527 | -0.00089 | -0.00003 |
| **Iowa** | -0.20357 | 0.27254 | -0.00047 | -0.00002 |
| **Mexico City** | -0.40187 | 0.43137 | -0.00013 | -0.00007 |
| **Malaysia** | -0.05247 | 0.04792 | -0.00372 | 0.00040 |
| | **SQNR** | | | |
| **Antarctica** | -0.17290 | -0.49905 | -0.00428 | -0.00018 |
| **Brazil** | -0.15764 | -0.03781 | -0.00115 | -0.00014 |
| **Greenland** | 0.03915 | -0.15997 | 0.00111 | 0.00003 |
| **Iowa** | 0.11695 | -0.33340 | 0.00094 | 0.00004 |
| **Mexico City** | 0.24756 | -0.19946 | -0.00148 | -0.00000 |
| **Malaysia** | 0.01359 | 0.00774 | -0.00420 | -0.00018 |
| | **Phase Difference** | | | |
| **Antarctica** | 0.06301 | 0.33723 | 0.01376 | 0.00126 |
| **Brazil** | 0.05659 | 0.04052 | -0.00192 | -0.00041 |
| **Greenland** | 0.03134 | 0.21307 | -0.01814 | -0.00097 |
| **Iowa** | -0.03299 | 0.38846 | -0.00639 | -0.00106 |
| **Mexico City** | -0.23009 | 0.20473 | -0.00496 | -0.00180 |
| **Malaysia** | -0.00571 | 0.02877 | 0.00309 | -0.00000 |
| | **Coherence Loss$_{\%}$** | | | |
| **Antarctica** | 0.23249 | 0.54473 | 0.01171 | 0.00062 |
| **Brazil** | 0.15347 | 0.01422 | 0.00009 | 0.00030 |
| **Greenland** | 0.02646 | -0.02993 | -0.00050 | -0.00053 |
| **Iowa** | -0.00192 | 0.38737 | -0.00699 | -0.00129 |
| **Mexico City** | -0.01596 | -0.06535 | 0.00209 | -0.00154 |
| **Malaysia** | -0.01573 | 0.03333 | 0.00280 | -0.00007 |

Table 7: Signal performance correlation with locally computed statistics.

of linear correlation that might exist between locally computed image statistics and the performance loss introduced by the 8:2 BAQ compression scheme. The performance loss is evaluated on the focused images in terms of four different metrics:

- Quantization noise magnitude, computed as: $|\eta_{baq}| = |I_{adc} - I_{baq}|$;
- Signal-to-quantization-noise ratio, computed as: $SQNR = \frac{|I_{adc}|^2}{|I_{adc}-I_{baq}|^2}$;
- Absolute phase difference, computed as: $|\Delta\phi| = |\angle I_{adc} - \angle I_{baq}|$, wrapping values that fall outside the $[-\pi; +\pi]$ range;
- Coherence loss percentage, computed as: $\Delta\gamma_{\%} = \frac{|\gamma_{adc}-\gamma_{baq}|}{|\gamma_{adc}|} \cdot 100$;

where $I_{adc}$ and $I_{baq}$ are the 8-bit and 2-bit SLC images respectively, and $\gamma_{adc}$ and $\gamma_{baq}$ are the correspondent interferometric coherence products as described in Section 2.3.4. A set of image features is computed for each acquisition. This consists of the following locally computed statistics:

- Average;
- Standard Deviation;
- Skewness;
- Kurtosis.

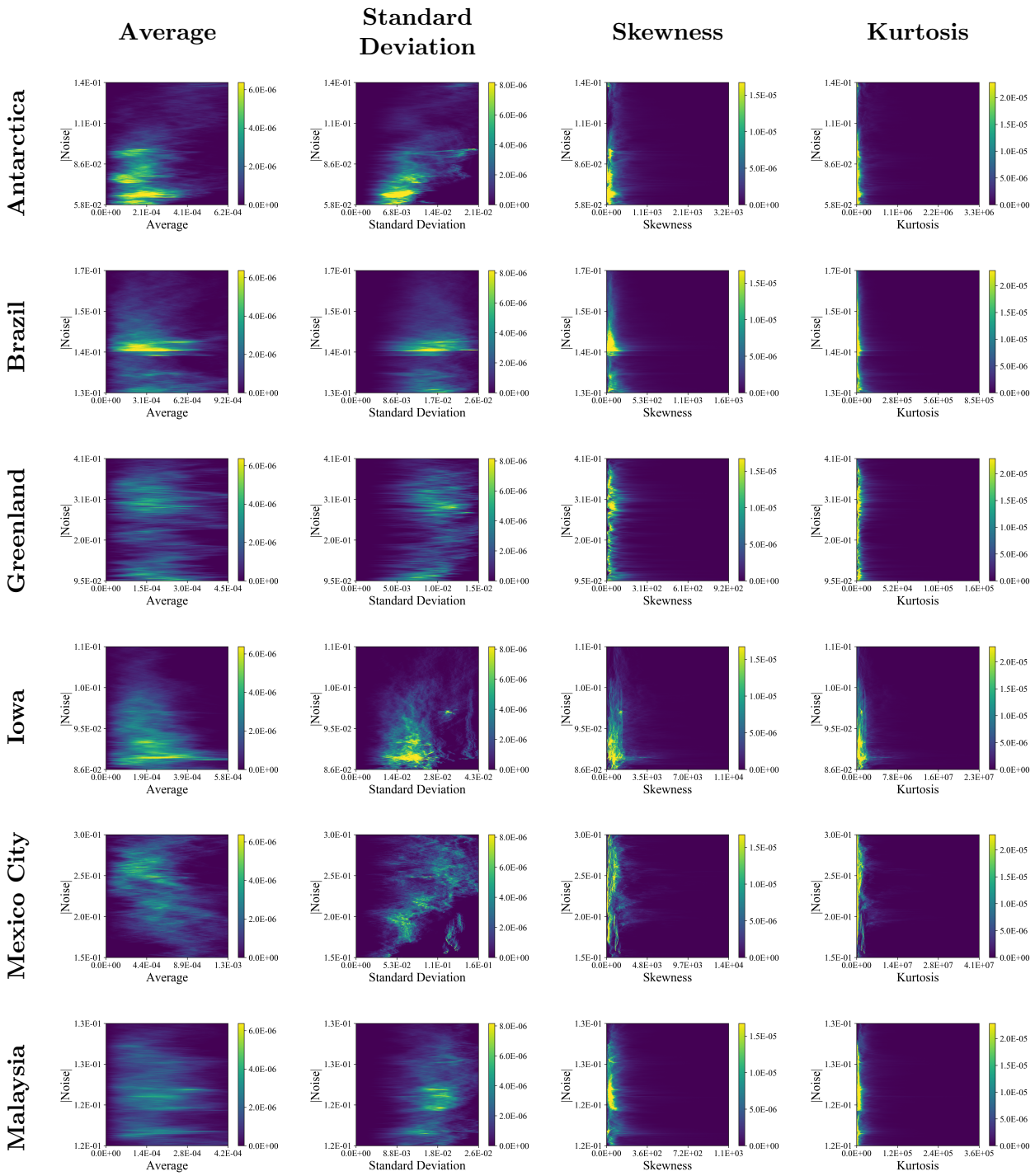



Figure 30: Absolute error scatterplots compared to locally computed statistics.

Each pixel value in the resulting feature-maps, consists of image statistics computed considering a sliding window of the previously computed range and azimuth dimensions, as described in Table 6. Considering the large dimensions of each acquisition, the sliding window process has been efficiently implemented in the Discrete Fourier Transform (DFT) domain using the Fast Fourier Transform (FFT) algorithm. Since the pixels on the borders of each image require the latter to be padded for the convolution process, potentially altering the true statistics of the underlying signal, all resulting feature- and performance-maps have been stripped off their border pixels by an amount equal to $0.5 \cdot (\mathcal{W} - 1)$, where $\mathcal{W}$ is a generic dimension of the sliding windows. Finally, the linear correlation

between the obtained feature- and performance-maps is computed in terms of Pearson Correlation Coefficient (PCC, or Pearson's R). Commonly represented with R, the PCC for two random variables $X$ and $Y$ is defined as

$$R_{X,Y} = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y} \tag{79}$$

where $\sigma_X$ and $\sigma_X$ are the standard deviations of $X$ and $Y$ respectively, and $cov(\cdot)$ is the covariance measure of the two variables. The resulting correlation coefficients are shown in Table 7. Firstly, it is possible to notice that both skewness and kurtosis statistics present low linear correlation with all tested performance metrics and across all datasets. The locally computed average feature-map displays low correlation values with all performance metrics except for the quantization error magnitude, for which both non negligible linear correlation is detected for both the Antarctica and the Mexico City datasets. Finally, the local standard deviation map displays the greatest linear correlation with all performance metrics, with peaks of up to $R = 0.59$ with the quantization error magnitude and the Antarctica dataset. This result is somewhat expected, as the theory discussed in Section 2.3 concluded that variability in the backscatterer distribution is a major cause of signal dependant quantization noise, since phenomena like low scatter suppression become less negligible as the local variability increases. Consistent with this assumption, the acquisitions that show the highest linear correlations are those with the highest dynamic ranges, namely Antarctica, Mexico City and Iowa. Similarly, the negative correlation between standard deviation and SQNR is also justifiable since a high backscatter variability causes performance degradation, due to the higher impact of signal-dependent quantization noise contributions.

In conclusion, the standard deviation metric showed the highest potential for qualitative performance estimation. Despite this, none of the testes moments achieved high enough correlation with the chosen performance metrics to allow for their direct quantification It is therefore not possible to give an estimate of the potentially achievable data-rate reduction.

# 4 Proposed CNN Denoising

In the following chapter, the potentials of deep learning for quantization noise removal in SAR images are investigated. Starting from an overview of the utilized quantization noise model, the proposed denoising architecture, the training set generation pipeline, and the training strategy are presented. Finally, performance metrics for the resulting architecture are discussed and compared to the traditional methods analyzed in Section 3.1.

## 4.1 Concept

Remote sensing SAR missions, especially spaceborne ones, are limited in the amount of acquirable data by the on-board memory capacity, and the downlink bandwidth and availability. The overall memory constraints are challenged by the rising data volume requirements associated with higher bandwidth usage, increased pulse repetition frequencies, and multiple polarizations of the most recent SAR systems. These limits are feasibly surmountable by using larger downlink bandwidths and by allocating more ground stations for increased coverage, at the cost of inflating the operational costs, which are often hardly justifiable, especially for systems close to the end of their planned lifetime. To reduce the pressure onto the memory, lower bit-rates could be used during quantization, with the undesirable drawback of lower signal-to-noise ratios. Similarly, the implementation of more sophisticated quantization techniques is limited by the computational resources available on-board. The proposed deep learning architecture discussed in this chapter explores the possibility of combining a simple, low bit-rate, on-board compression scheme with an on-ground image denoising mechanism. Such a processing pipeline could potentially reduce the memory and infrastructure requirements for on-going missions, while partially recouping some of the lost performance using a deep-learning-based denoiser.

## 4.2 Quantization Noise Model

Denoising is a crucial process for many applications pertaining to information theory. We discussed in Section 2.4.5 that for natural images, a clean signal $\boldsymbol{x}$ and a noisy observation $\boldsymbol{y}$ are related according to equation (73), where $\boldsymbol{\eta}$ represents the noise contribution present in $\boldsymbol{y}$. Typically, the noise figure $\boldsymbol{\eta}$ is assumed to be additive white Gaussian noise (AWGN) with standard deviation $\boldsymbol{\sigma_\eta}$. This assumption is satisfying for many natural stochastic processes. Conveniently, it is also a typical assumption used to model quantization noise, in that it is virtually uncorrelated with the affected signal when a large enough bit-rate is used [6]. This statement does not imply that the additive noise behaviour is always a valid assumption. In Section 2.3, the most common noise sources related to the quantization process have been explored. While granular and overload noises can be modelled as an additive noise process, low amplitude suppression and low scatter suppression are non-linear, signal-dependent sources of error, defying the hypothesis of additive independent Gaussian noise for quantization errors [13]. Since in the case of TSX- and TDX-acquired raw data, the granular noise can be considered to be the dominant quantization error component [12], for the sake of this work, the quantization noise will be modelled as an additive process. Crucially, speckle represents another noise component intrinsic of the SAR acquisition process, which will not be targeted by the proposed denoising process. Thus, it is important to underline that a deep learning architecture would be required to learn the distinction between speckle and quantization noise to avoid introducing unwanted distortions. In the following, to mitigate the added complexity introduced by the speckle phenomenon, a specific training strategy is proposed and discussed.

## 4.3 Proposed CNN Architecture

The proposed CNN denoising architecture is based on the denoising convolutional neural network (DnCNN) presented in [22]. It consists of a feedforward CNN, specifically adopting a deep residual learning architecture and batch-normalization layers for faster training and better performance. Notably, this design does not rely on pooling layers. Thus the receptive field ($RF$) for a network of depth $D$, using convolutional kernels of size $K \times K$ pixels, is given by

$$RF = (K-1)D + 1 \tag{80}$$

Typically, larger amounts of noise require a larger receptive field to capture more spatial information for image restoration. For this application, a kernel size $K$ of $3 \times 3$ pixels has been used for each of the 17 network layers, resulting in a receptive field of $35 \times 35$ pixels. The input, output, and hidden layer utilize three different layer arrangements:

- Conv+ReLU: the input layer consists of a convolutional layer using 32 filters of size $3 \times 3 \times C$ to produce 32 features with $C$ channels each; this is followed by a rectified linear unit (ReLU), utilized to introduce nonlinearities in the architecture;
- Conv+BN+ReLU: the following $(D-2)$ hidden layers each consists of a convolutional layer using 32 kernels of shape $3 \times 3 \times C$, followed by a batch normalization step before the final ReLU activation function;
- Conv: the final layer consists of a convolutional layer using a single $3 \times 3 \times 32$ pixel-sized filter to reconstruct the output.

To reduce boundary artifacts, each convolutional layer adopts a pre-processing step consisting of the symmetrical padding of the input patch before the convolution step. Given the noisy input to the network expressed in (73), the network is set-up to learn the residual mapping function $\mathcal{F}(\boldsymbol{y}) = \hat{\boldsymbol{\eta}}$, such that

$$\hat{\boldsymbol{x}} = \boldsymbol{y} - \mathcal{F}(\boldsymbol{y}) \tag{81}$$

where $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{\eta}}$ are used to denote the predicted values of $\boldsymbol{x}$ and $\boldsymbol{\eta}$, respectively. Thus an identity shortcut is used to bypass the input image to the output.

## 4.4 Training Set Generation

The dataset utilized to generate the training-set consists of the six TanDEM-X acquisitions described in Section 3.3.2. The entire training-set generation consists of a parametric pipeline, intelligently extracting samples depending on the user-specified target statistics. All tests have been performed considering the intensities of the acquisitions, thus all input single-look complex images have been pre-processed accordingly. Each image is then sampled with a stride length of 32 pixels into patches of $64 \times 64$ pixels, resulting in an oversampled representation of the input dataset. This data augmentation process has been accounted for by vertically and horizontally flipping each second patch, paying attention to avoid $\pm\pi/2$ patch rotations. Indeed, azimuth and range directions should be preserved to avoid alterations in the underlying noise model. For this reason, only left-right and top-bottom flips have been considered to diversify the data set. Then, for each extracted patch, the combined pixel statistics are computed together with the corresponding quantization noise magnitude feature. The latter is used to cluster the resulting patch noise magnitude distribution into eight equally spaced bins, plus two specific bins used to cluster the first and the last $10^{th}$ percentiles (i.e., the tails) of the distribution.

The resulting bins are used to select the available patches according to the desired target distribution of the training-set, with two particular possibilities:

- Homogeneous sampling, meaning that for each bin, the same number of patches is extracted. It results in an equally represented sample population, at the cost of a smaller-sized training-set, since the less populated bin limits the overall number of samples.

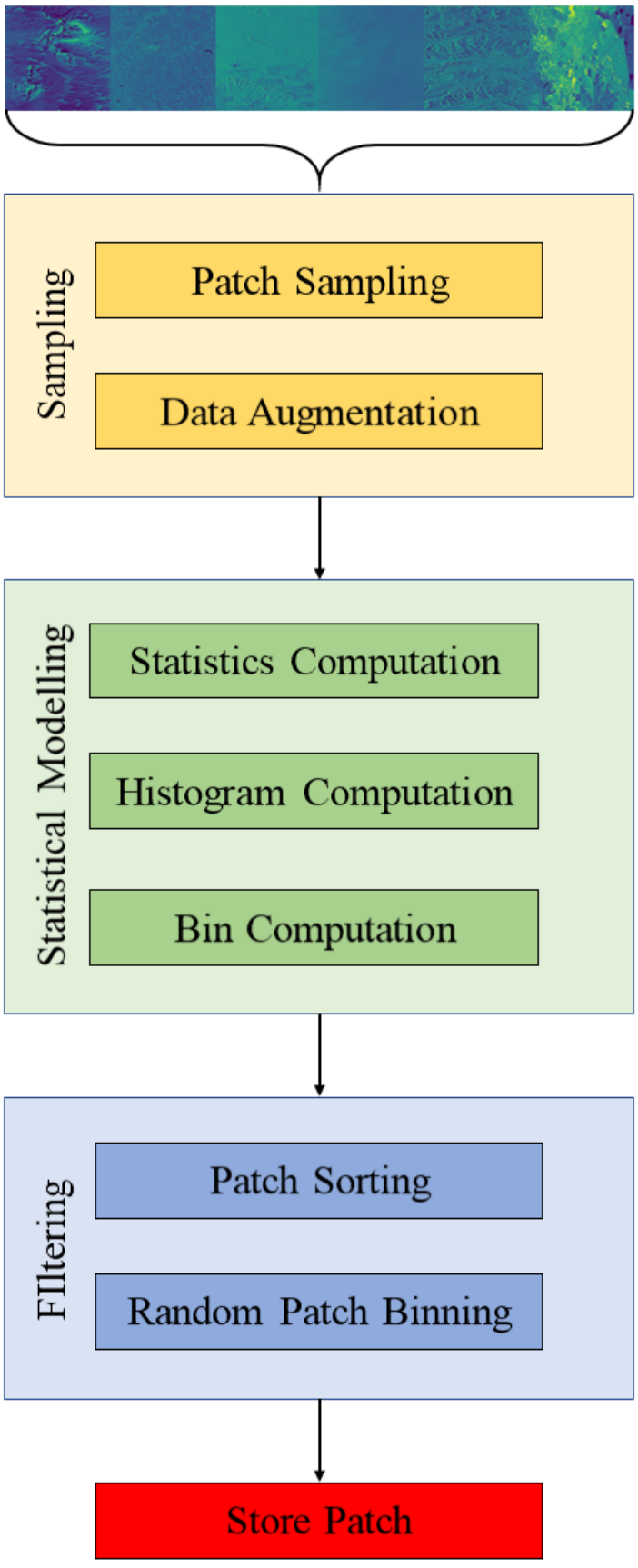



Figure 31: Training-set generation pipeline.

- Heterogeneous sampling, meaning that for each bin, a fraction of the available samplesa is selected. Such a sampling strategy results in a training-set distribution equal to that of the original dataset, allowing for the maximum training-set size.

Thus, the chosen strategy determines the number of samples-per-bin composing the training-set, which are then chosen randomly to properly mix patches belonging to different acquisitions. A 10% fraction of the available samples-per-bin is kept aside to build an unbiased test-set to measure the network's performance. In contrast, 10% of the sampled training-set constitutes the validation-set used during training to spot anomalies like over- or under-fitting. The selected patches are extracted from their corresponding acquisitions and stored individually, allowing to flexibly pick a mini-batch size at run-time without needing to recompute the training-set.

## 4.5 Training Strategy

As anticipated in Section 4.2, an ad hoc training solution has been devised to ignore the adverse effects associated with the speckle phenomena. Instead of directly computing the loss function on the denoised image, as it is produced by the architecture described in Section 4.3, first a despeckling filter is used to mitigate the influence of the corresponding phenomena, then the loss function is computed on the resulting image, thus effectively bypassing the variability introduced by the SAR speckle.

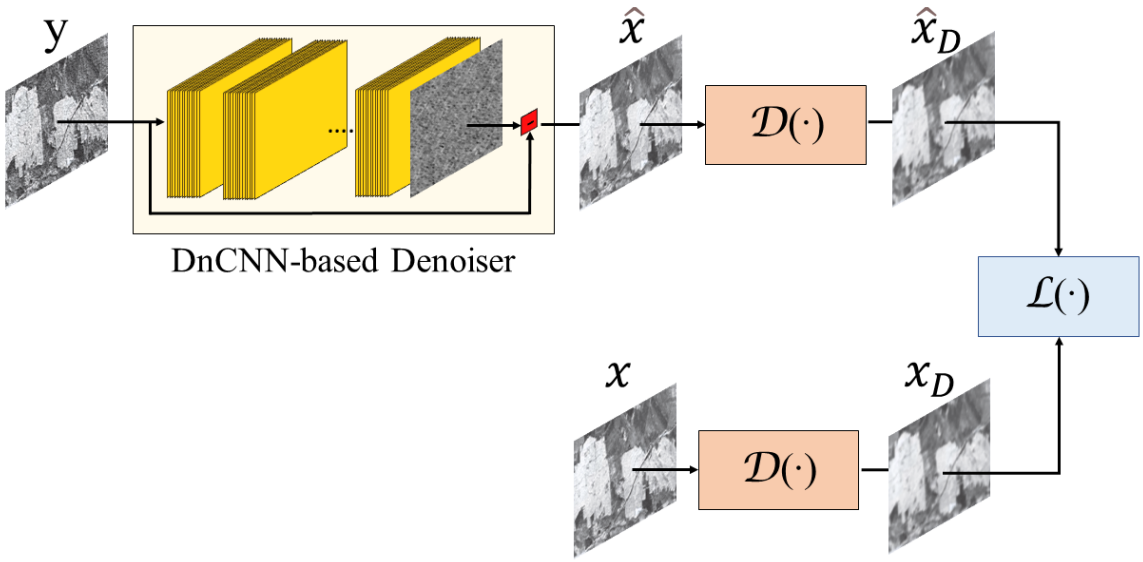



Figure 32: Training strategy for the proposed denoiser. The CNN generates its prediction from the noisy input image $y$. The resulting denoised image $\hat{x}$ is passed through a despeckling filter $\mathcal{D}(\cdot)$. The resulting image $\hat{x}_{\mathcal{D}}$ is compared to the despeckled target image $x_{\mathcal{D}}$ to compute the loss function $\mathcal{L}(\cdot)$.

A simple $3 \times 3$ boxcar averaging-filter (boxcar) has been chosen for this purpose, but other more sophisticated solutions could be used instead. Effectively, this despeckling step could be seen as an additional layer added in cascade to the proposed network. It was in the interest of this thesis' work to assess and compare the performance gains obtained by the network before and after the boxcar filter, therefore the denoising architecture and the despeckling filter have been kept separated. The resulting loss function, used to train the denoising architecture, effectively consists of the mean squared error (MSE) metric of the $3 \times 3$ averaged output with respect to the $3 \times 3$ averaged target. After extensive testing, it was decided to stop all training phases after 30 epochs. The reasoning being that for most of the tests discussed in this work, the training losses indicated convergence after 20 to 30 epochs. On the other hand, the experiments adopting smaller-sized training-sets showed a substantial increase in the generalization gap defined by the training and the validation losses after 30 epochs.

## 4.6 Experiments and Results

In the following section, the denoising performance of the proposed network is evaluated for different types of training-sets. All tests rely on the MSE metric, which is used during the training phase as a loss function, and on the SQNR metric, which is used to assess the overall performance of the filtered patches.

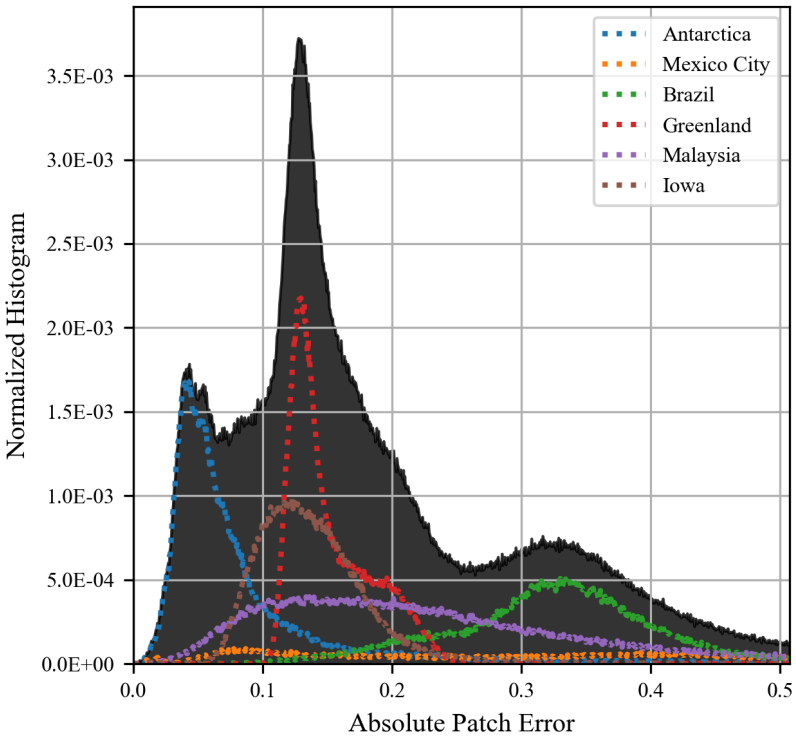



Figure 33: Average patch error composition for the combined dataset. The contribution of each individual acquisition is plotted as a colored dotted line.

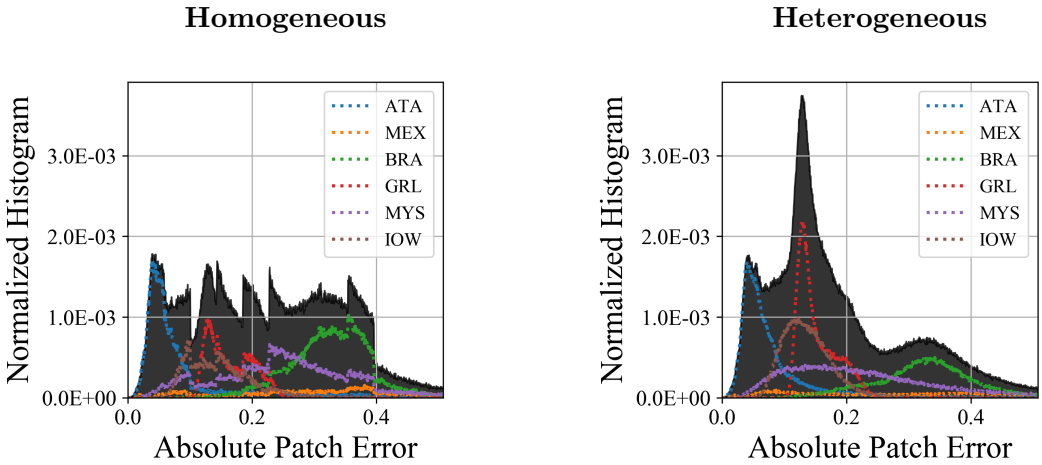



Figure 34: Training-set composition after using a homogeneous (left) and heterogeneous sampling strategy (right). The contribution of each individual acquisition to the final distributions is plotted as a dotted line.

### 4.6.1 Combined Dataset

As a first experiment, all six acquisitions composing the dataset described in Section 3.3.2 have been considered. Following the procedure explained in Section 4.4, the resulting absolute patch error distribution for the combined dataset is shown in Figure 33.

| | **Homogeneous** | **Heterogeneous** |
|---|---|---|
| **Combined** | 739351 | 1693211 |

Table 8: Sample-count comparison for the combined datasets using homogeneous and heterogeneous sampling strategies .

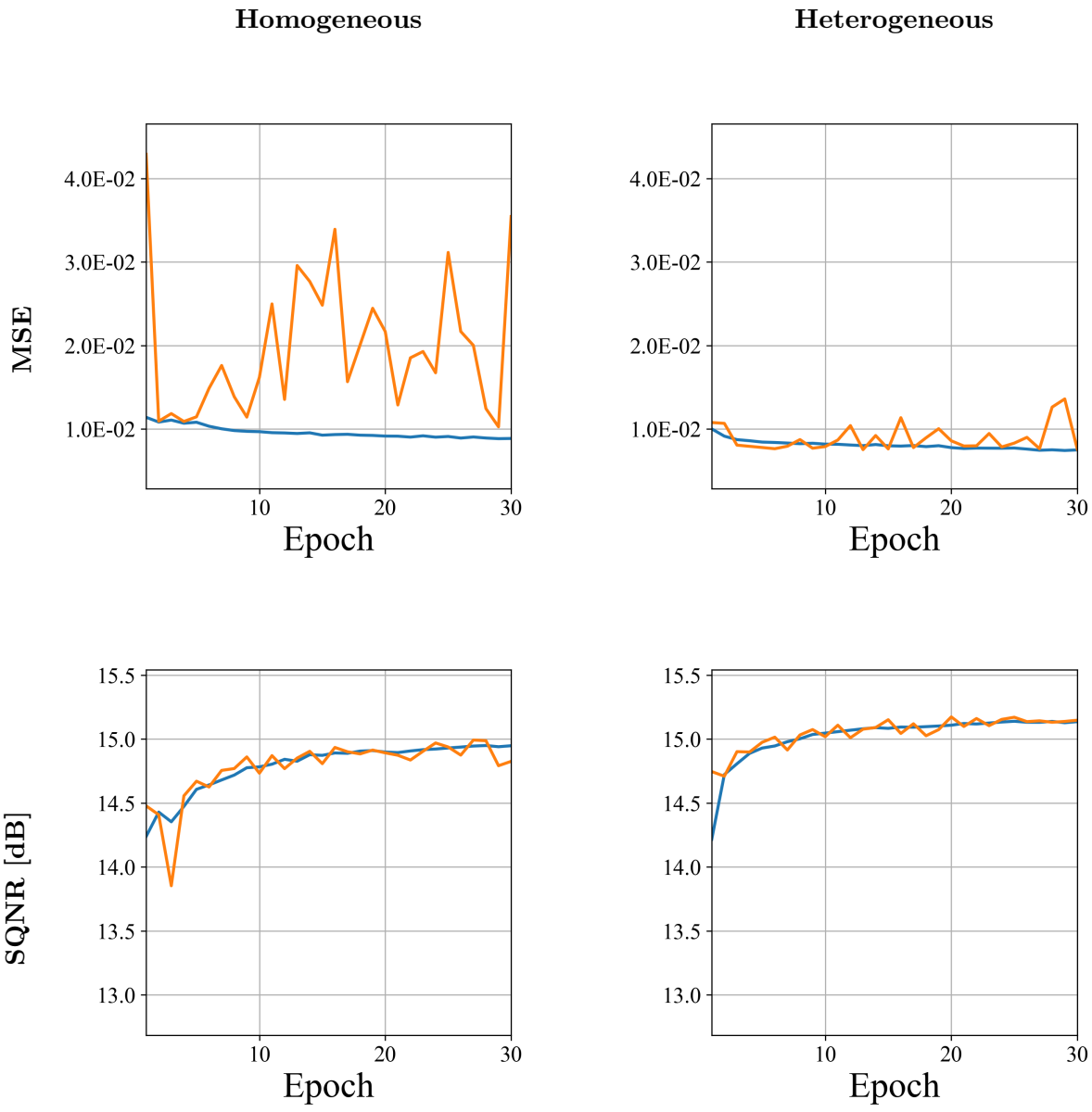



Figure 35: Training metric (in blue) and validation metric (in orange) evolutions for the combined dataset. The training loss function is represented in the first row, while the SQNR metric has been plotted in the second one. The first row represents the evolution using a homogeneously-sampled dataset, while the right one represents that resulting from a heterogeneously-sampled one.

The normalized histogram, plotted in black, shows the overall patch error distribution for the combined dataset, while the normalized contributions of each acquisition are plotted using individual colors. It is possible to notice that all images are unique and constitute highly varying distributions. In particular, the Mexico City and the Malaysia acquisitions are associated with very broad distributions, suggesting that the high variability observed in Section 3.3 is effectively reflected in the large absolute patch error variance displayed by the histogram. Following this realization, the datasets resulting from both the heterogeneous and the homogeneous sampling strategies have been compared to each other, testing whether such differences in land covers and associated quantization error distribution can be mitigated by uniformly sampling patches across the overall distribution. The two normalized histograms associated with the resulting training-sets are displayed in Figure 34. It is possible to notice that the homogeneous sampling strategy produces a uniformly distributed dataset. The only exception is the tail of the distribution (absolute patch error values greater than 0.4), which contains the last $10^{th}$ percentile of the original distribution. The contributions of individual acquisitions are also shown to be more uniform. As expected, the distribution resulting from the heterogeneous sampling strategy mirrors that of the complete dataset. Following these sampling strategies, the resulting training-set sample counts are displayed in Table 8. These numbers confirm the reduced sample count produced by the homogeneous sampling strategy, as the sample-per-bin count is limited by the bin associated with the smallest population. The proposed network possesses 139.808 trainable parameters. Thus both

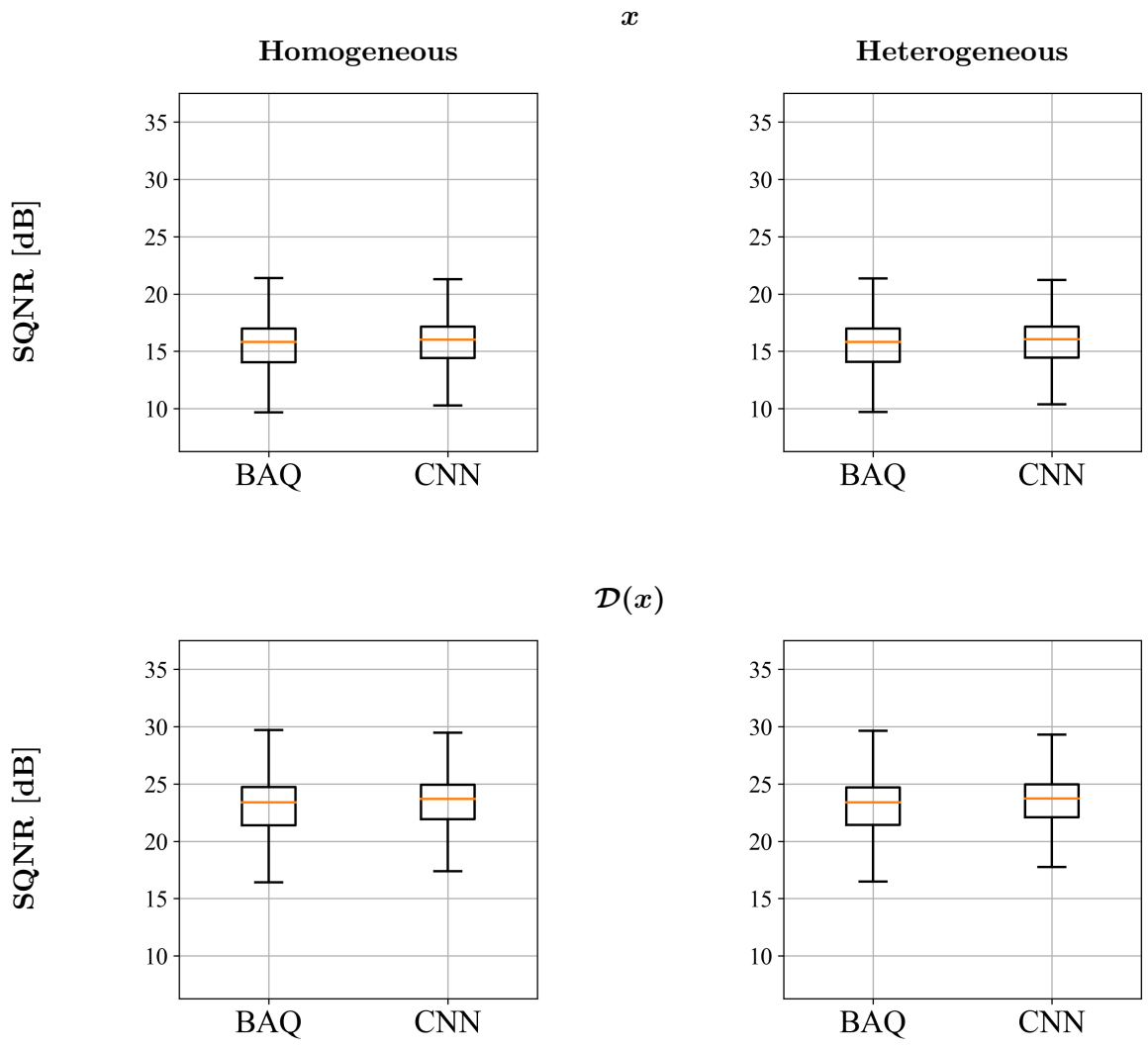



Figure 36: Boxplots of the performance in terms of SQNR for the combined training-set.

| | $MSE_{GAIN}$ | $SQNR_{GAIN}$ | $MSE_{GAIN}$ | $SQNR_{GAIN}$ |
|---|---|---|---|---|
| | $x$ | | $\mathcal{D}(x)$ | |
| **Homogeneous** | $4.32 \cdot 10^{-3}$ | 0.26 | $1.68 \cdot 10^{-3}$ | 0.43 |
| **Heterogeneous** | $6.15 \cdot 10^{-3}$ | 0.31 | $2.01 \cdot 10^{-2}$ | 0.54 |

Table 9: Average performance gains for the combined training-set.

training-sets should be adequate to train the architecture properly. The training and validation losses for both training-sets are displayed in Figure 35, together with the corresponding SQNR evolutions.

When the training-losses are considered, both networks show a rapid stagnation in their evolution. However, the validation losses suggest that the network associated with the homogeneous training-set had difficulties converging, with a large and erratic generalization gap appearing after the fourth epoch. On the other hand, the heterogeneous network shows better conversion covergence, with a smaller generalization gap appearing after the $27^{th}$ epoch. The SQNR evolution showed an improvement during the training, with the heterogeneously-trained network concluding with a 0.2 dB performance gain over the homogeneous one. Both networks have been evaluated using the same unbiased test-set to measure their relative performance gains over the noisy BAQ patches. For this purpose, the mean squared error and the signal-to-quantization-noise ratio metrics have been used to evaluate the patches both before the despeckling filter ($x$) and after the filter ($\mathcal{D}(x)$). The resulting performance gains are displayed in Table 9, with positive values corresponding to performance improvement. It is possible to notice that the despeckled patches showed more significant performance gains for both training-sets than those that have not been filtered. As the training losses suggested, the network associated with the heterogeneous training-set performed better. The box-plots associated with the test-sets' signal-to-quantization-noise ratios are shown in Figure 36. These graphs suggest that the networks saw the most significant improvements for those patches with lower SQNR.

### 4.6.2 Individual Datasets

To test whether the varied training-set, described in Section 4.6.1, might be too challenging to model the proposed network architecture, the tests presented in the previous section have been repeated for each acquisition, separately. Sampling using both homogeneous and heterogeneous strategies, yields the sample count displayed in the Table 10. All datasets, except for Mexico City and Antarctica, posses comparable numbers of samples to the number of trainable parameters in the deep neural network. On the contrary, Mexico City is significantly deficient in its sample population, as it also possesses the highest sample variance. The respective training losses are plotted in Figure 37. Interestingly, the training losses for Brazil, Greenland, Iowa, and Antarctica datasets all appear to converge, despite the Antarctica one, mainly being made-up of relatively few samples. The respective corresponding SQNR evolutions show only little performance gain after the $20^{th}$ epoch, again suggesting that the associated networks have successfully converged during training. On the other hand, Mexico City and Malaysia show vastly erratic validation losses, meaning that the networks had trouble learning the underlying noise patterns from the training-sets. In particular, the behavior of the Mexico City dataset appears to be linked to the small sample count, as the generalization gap grows as the training loss decreases,

| | **Homogeneous** | **Heterogeneous** |
|---|---|---|
| **Antarctica** | 49391 | 334496 |
| **Brazil** | 144331 | 266773 |
| **Greenland** | 160251 | 318843 |
| **Iowa** | 142051 | 298536 |
| **Mexico City** | 24701 | 126713 |
| **Malaysia** | 136771 | 347806 |

Table 10: Sample-count comparison for the individual datasets using homogeneous and heterogeneous sampling strategies .

hinting at potential overfitting issues over the small training-set. Evaluating the trained networks yields the average performance gains reported in the table 11. These metrics show that, on average, when compared to the combined dataset results, the performance gains are more significant for all individual datasets except for Iowa, with the networks associated with the Mexico City and Malaysia training-set showing the most significant improvements. This unexpected result might be related to the overall lower signal-to-quantization-noise ratios of these test-sites, meaning that the most significant gains are associated to those acquisitions showing the most deficient starting performance.

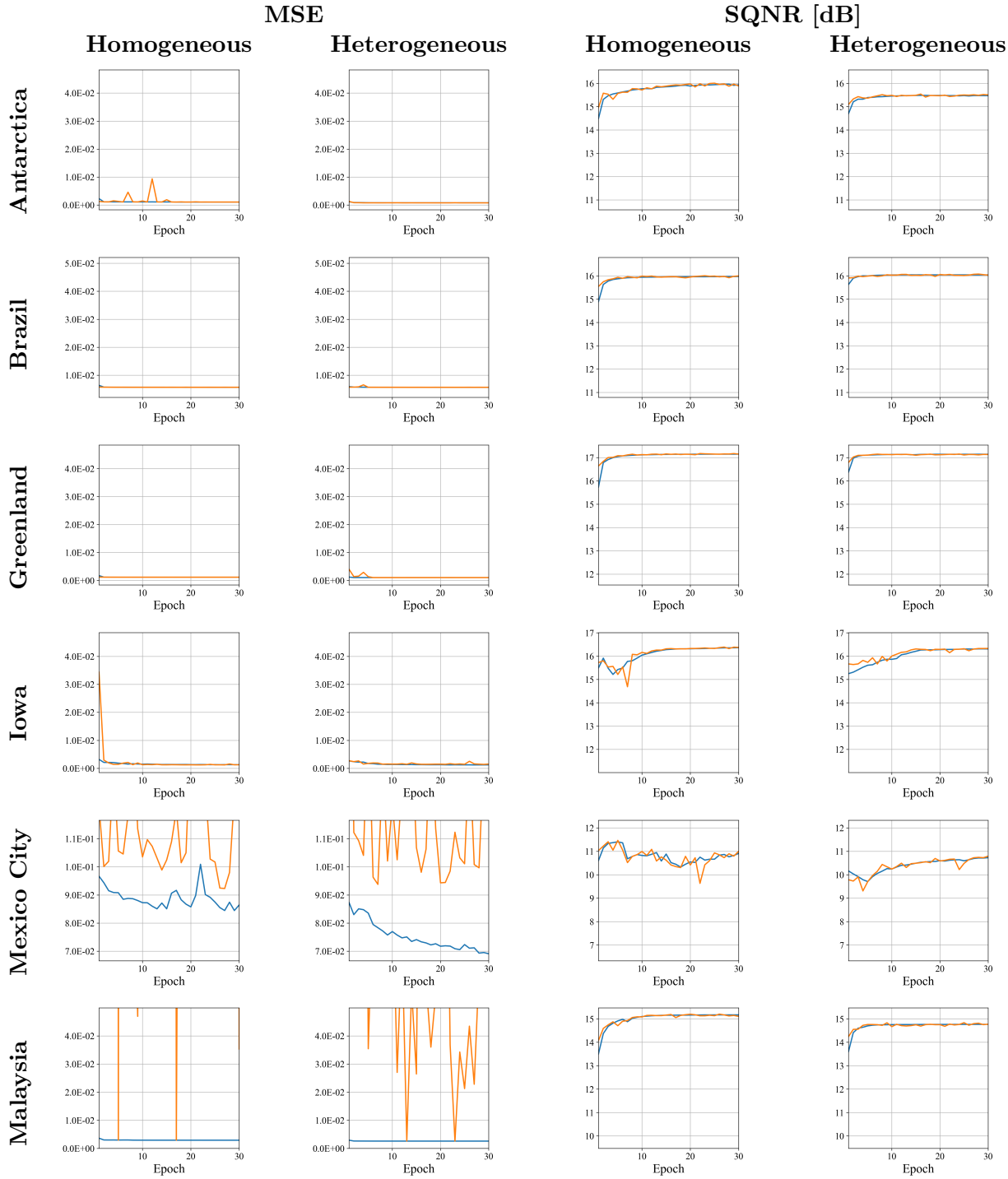



Figure 37: Training loss evolution (in blue) and validation loss evolution (in orange) for the individual datasets.

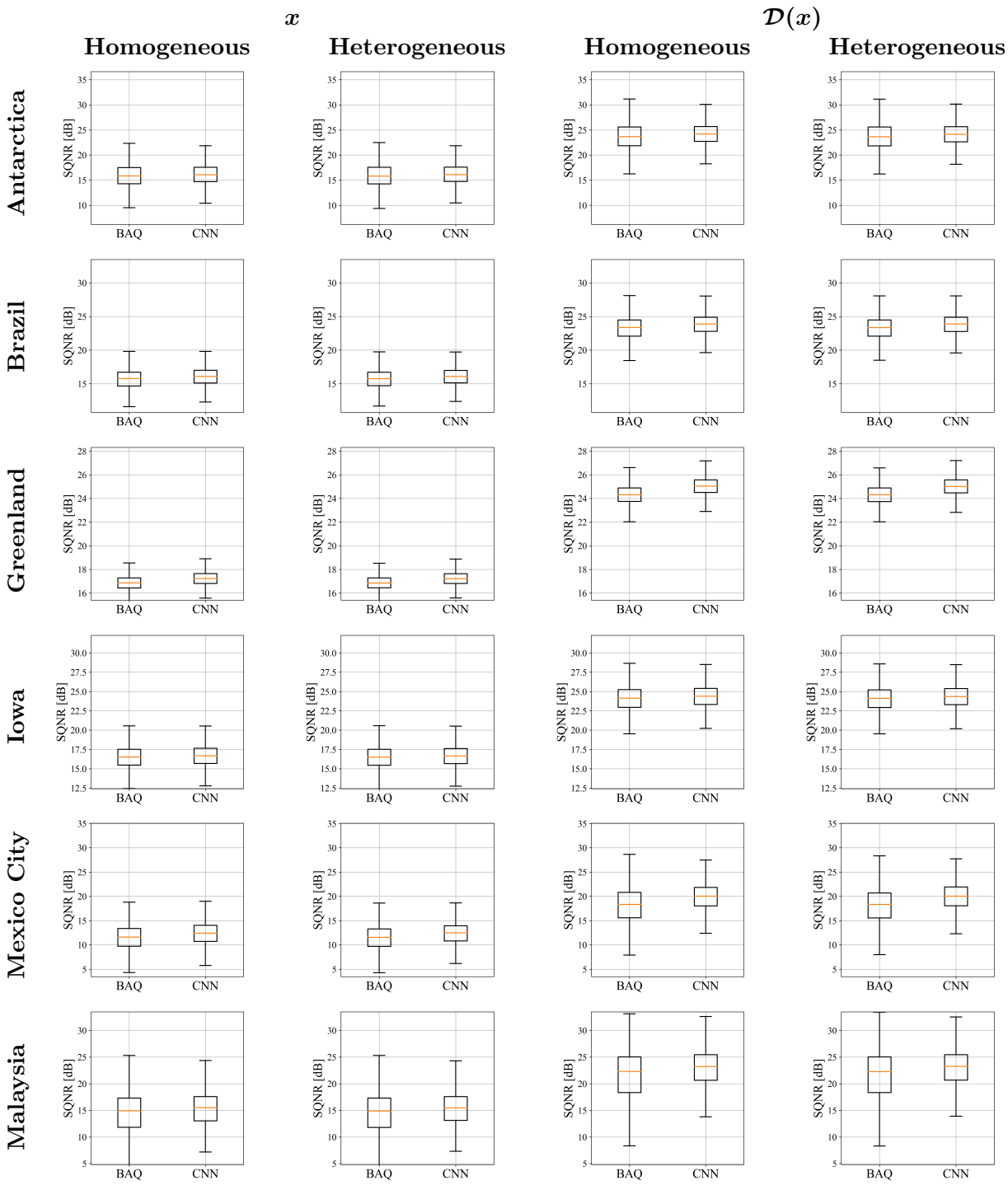



Figure 38: SQNR boxplots for the individual test-sets.

.

These metrics show that, on average, when compared to the combined dataset results, the performance gains are more significant for all individual datasets except for Iowa, with the networks associated with the Mexico City and Malaysia training-set showing the most significant improvements. This unexpected result might be related to the overall lower signal-to-quantization-noise ratios of these test-sites, meaning that the most significant gains are associated to those acquisitions showing the most deficient starting performance. All networks see their most significant performance uplift when the despeckled patches are tested, replicating the combined dataset findings. Finally, all heterogeneously-

| | | $MSE_{GAIN}$ | $SQNR_{GAIN}$ | $MSE_{GAIN}$ | $SQNR_{GAIN}$ |
|---|---|---|---|---|---|
| | | $\boldsymbol{x}$ | | $\mathcal{D}(\boldsymbol{x})$ | |
| **ANT** | **Homogeneous** | $3.52 \cdot 10^{-4}$ | 0.28 | $1.24 \cdot 10^{-4}$ | 0.53 |
| | **Heterogeneous** | $3.86 \cdot 10^{-4}$ | 0.32 | $1.40 \cdot 10^{-4}$ | 0.53 |
| | | $\boldsymbol{x}$ | | $\mathcal{D}(\boldsymbol{x})$ | |
| **BRA** | **Homogeneous** | $1.63 \cdot 10^{-3}$ | 0.38 | $5.57 \cdot 10^{-4}7$ | 0.61 |
| | **Heterogeneous** | $1.66 \cdot 10^{-3}$ | 0.37 | $5.61 \cdot 10^{-4}$ | 0.61 |
| | | $\boldsymbol{x}$ | | $\mathcal{D}(\boldsymbol{x})$ | |
| **GRL** | **Homogeneous** | $7.19 \cdot 10^{-4}$ | 0.37 | $2.83 \cdot 10^{-4}$ | 0.73 |
| | **Heterogeneous** | $7.19 \cdot 10^{-4}$ | 0.36 | $2.79 \cdot 10^{-4}$ | 0.71 |
| | | $\boldsymbol{x}$ | | $\mathcal{D}(\boldsymbol{x})$ | |
| **IOW** | **Homogeneous** | $2.46 \cdot 10^{-3}$ | 0.18 | $1.11 \cdot 10^{-3}$ | 0.29 |
| | **Heterogeneous** | $2.44 \cdot 10^{-3}$ | 0.17 | $1.12 \cdot 10^{-3}$ | 0.30 |
| | | $\boldsymbol{x}$ | | $\mathcal{D}(\boldsymbol{x})$ | |
| **MEX** | **Homogeneous** | $4.44 \cdot 10^{-2}$ | 0.83 | $7.84 \cdot 10^{-3}$ | 1.785 |
| | **Heterogeneous** | $4.22 \cdot 10^{-2}$ | 0.94 | $2.87 \cdot 10^{-2}$ | 1.971 |
| | | $\boldsymbol{x}$ | | $\mathcal{D}(\boldsymbol{x})$ | |
| **MYS** | **Homogeneous** | $5.87 \cdot 10^{-4}$ | 0.88 | $2.47 \cdot 10^{-4}$ | 1.62 |
| | **Heterogeneous** | $6.31 \cdot 10^{-4}$ | 0.93 | $2.69 \cdot 10^{-4}$ | 1.67 |

Table 11: Average performance gain for the individual training-sets.

sampled training-sets see comparable or noticeably greater performance gains when compared to the homogeneously-sampled ones, hinting again at lack of samples for some of the test-sites. Figure 38 shows the boxplots of the SQNR metrics computed on the test-sets.

All boxplots suggest that the most significant performance uplifts correspond to the worst-performing patches in terms of SQNR. It is also worth noticing that the Malaysia and Mexico City networks display the greatest performance dispersion, as all scatterplots are plotted in the same value range. In Figure

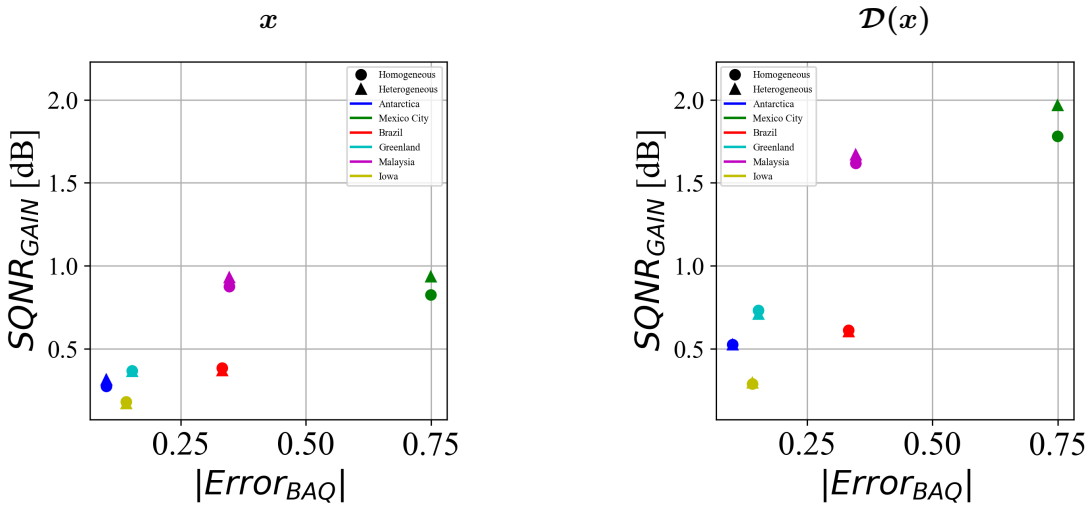



Figure 39: Average performance gain versus average quantization error magnitude, using both homogeneously- and a heterogeneously-sampled training-sets.

| | 1 Feature | 2 Features |
|---|---|---|
| **Antarctica** | 334496 | 196101 |
| **Brazil** | 266773 | 134204 |
| **Greenland** | 318843 | 193073 |
| **Iowa** | 298536 | 163165 |
| **Mexico City** | 126713 | 66646 |
| **Malaysia** | 347806 | 205059 |

Table 12: Sample-count comparison for the individual datasets using 1 and 2 input features.

39, the average SQNR performance-gains are reported for the individual dataset, distinguishing between heterogeneously- and homogeneously-sampled ones. The graphs show that greater performance gains are associated with those test-sites with the largest average quantization error magnitudes. The Malaysia dataset is a less obvious fit. On average, the error magnitude is comparatively low, but as the boxplots in Figure 38 show, this behaviour is explainable with its larger performance variance, with a significant sample population displaying considerably lower than average performances.

### 4.6.3 Footprint Feature

In Section 3.3, the sample neighborhood impacting the local quantization error has been explored, with the results reported in Table 6, which highlight the extents of these regions. It would be technically unfeasible to work with network architectures with receptive fields of similar dimensions, as the memory requirements for comparably deep networks would be prohibitive, and hundreds of acquisition would be necessary to extract plenty, large enough patches to train them. In order to mitigate this drawback and test whether the tested networks are missing some crucial global information to further improve

| | | $MSE_{GAIN}$ | $SQNR_{GAIN}$ | $MSE_{GAIN}$ | $SQNR_{GAIN}$ |
|---|---|---|---|---|---|
| **ANT** | | $\boldsymbol{x}$ | | $\boldsymbol{\mathcal{D}(x)}$ | |
| | **1 Feature** | $3.86 \cdot 10^{-4}$ | 0.32 | $1.40 \cdot 10^{-4}$ | 0.53 |
| | **2 Features** | $6.36 \cdot 10^{-4}$ | 0.40 | $2.59 \cdot 10^{-4}$ | 0.87 |
| **BRA** | | $\boldsymbol{x}$ | | $\boldsymbol{\mathcal{D}(x)}$ | |
| | **1 Feature** | $1.66 \cdot 10^{-3}$ | 0.37 | $5.61 \cdot 10^{-4}$ | 0.61 |
| | **2 Features** | $1.51 \cdot 10^{-4}$ | 0.42 | $5.66 \cdot 10^{-4}$ | 0.71 |
| **GRL** | | $\boldsymbol{x}$ | | $\boldsymbol{\mathcal{D}(x)}$ | |
| | **1 Feature** | $7.19 \cdot 10^{-4}$ | 0.36 | $2.79 \cdot 10^{-4}$ | 0.71 |
| | **2 Features** | $7.04 \cdot 10^{-4}$ | 0.41 | $2.92 \cdot 10^{-4}$ | 0.86 |
| **IOW** | | $\boldsymbol{x}$ | | $\boldsymbol{\mathcal{D}(x)}$ | |
| | **1 Feature** | $2.44 \cdot 10^{-3}$ | 0.17 | $1.12 \cdot 10^{-3}$ | 0.30 |
| | **2 Features** | $1.57 \cdot 10^{-3}$ | 0.20 | $5.98 \cdot 10^{-4}$ | 0.35 |
| **MEX** | | $\boldsymbol{x}$ | | $\boldsymbol{\mathcal{D}(x)}$ | |
| | **1 Feature** | $4.22 \cdot 10^{-2}$ | 0.94 | $2.87 \cdot 10^{-3}$ | 1.97 |
| | **2 Features** | $2.36 \cdot 10^{-3}$ | 1.22 | $2.43 \cdot 10^{-3}$ | 2.13 |
| **MYS** | | $\boldsymbol{x}$ | | $\boldsymbol{\mathcal{D}(x)}$ | |
| | **1 Feature** | $6.31 \cdot 10^{-4}$ | 0.93 | $2.69 \cdot 10^{-4}$ | 1.67 |
| | **2 Features** | $6.13 \cdot 10^{-4}$ | 1.03 | $2.64 \cdot 10^{-4}$ | 1.78 |

Table 13: Average performance gains for the individual training-sets, using one and two features respectively.

performance, a second input feature is added to complement the noisy signal intensity. This feature consists of the locally computed standard deviation map discussed in Section 3.3.3, as it possessed the highest average linear correlation with the quantization error magnitude. Since these feature maps are produced using large sliding windows, each acquisition's usable area is limited to the central portions. This aspect implies that the resulting training-sets will have significantly lower numbers of samples. To compensate for this deficiency, only heterogeneously-sampled training-set have been considered. The sample-counts for the training-sets, with one and two input features, respectively, are displayed

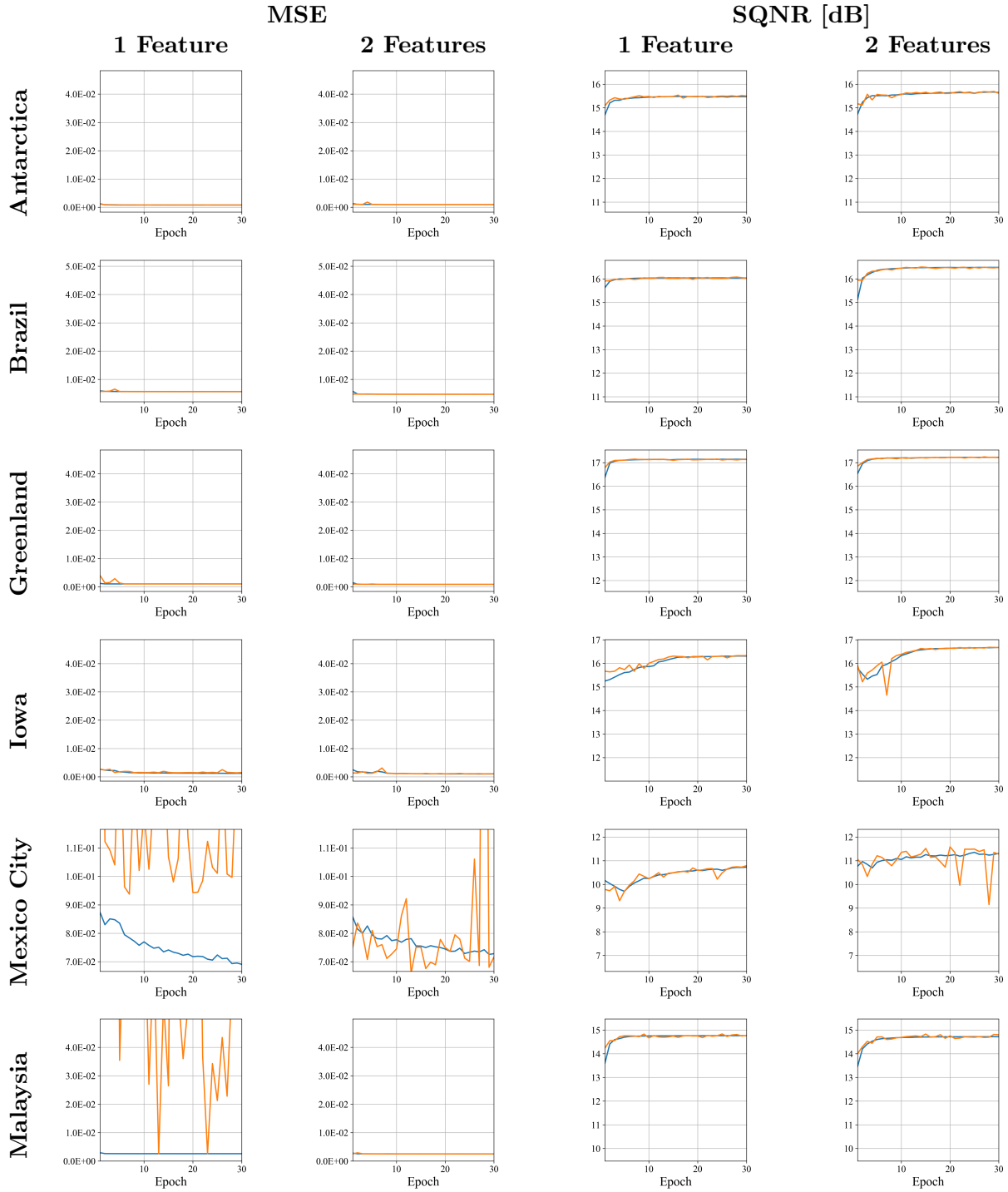



Figure 40: Training loss evolution (in blue) and validation loss evolution (in orange) for the networks associated to the individual datasets.

in Table 12. All training-sets are characterized by a comparable number of samples, with Mexico City being the only slightly underrepresented test-site. The respective training-loss evolutions are reported in Figure 40. It is of interest to notice that using two features, the Malaysia network converges. At the same time, the Mexico City case also significantly improves its validation loss, with a larger generalization gap appearing after 25 epochs. This suggests that the complementary feature has a positive impact on the tested networks' modelling capabilities. Performance gains for the trained networks are reported in Table 13. The results show that, the use of the proposed second feature allows for improving all networks performance gains, without any exception. This is also displayed in

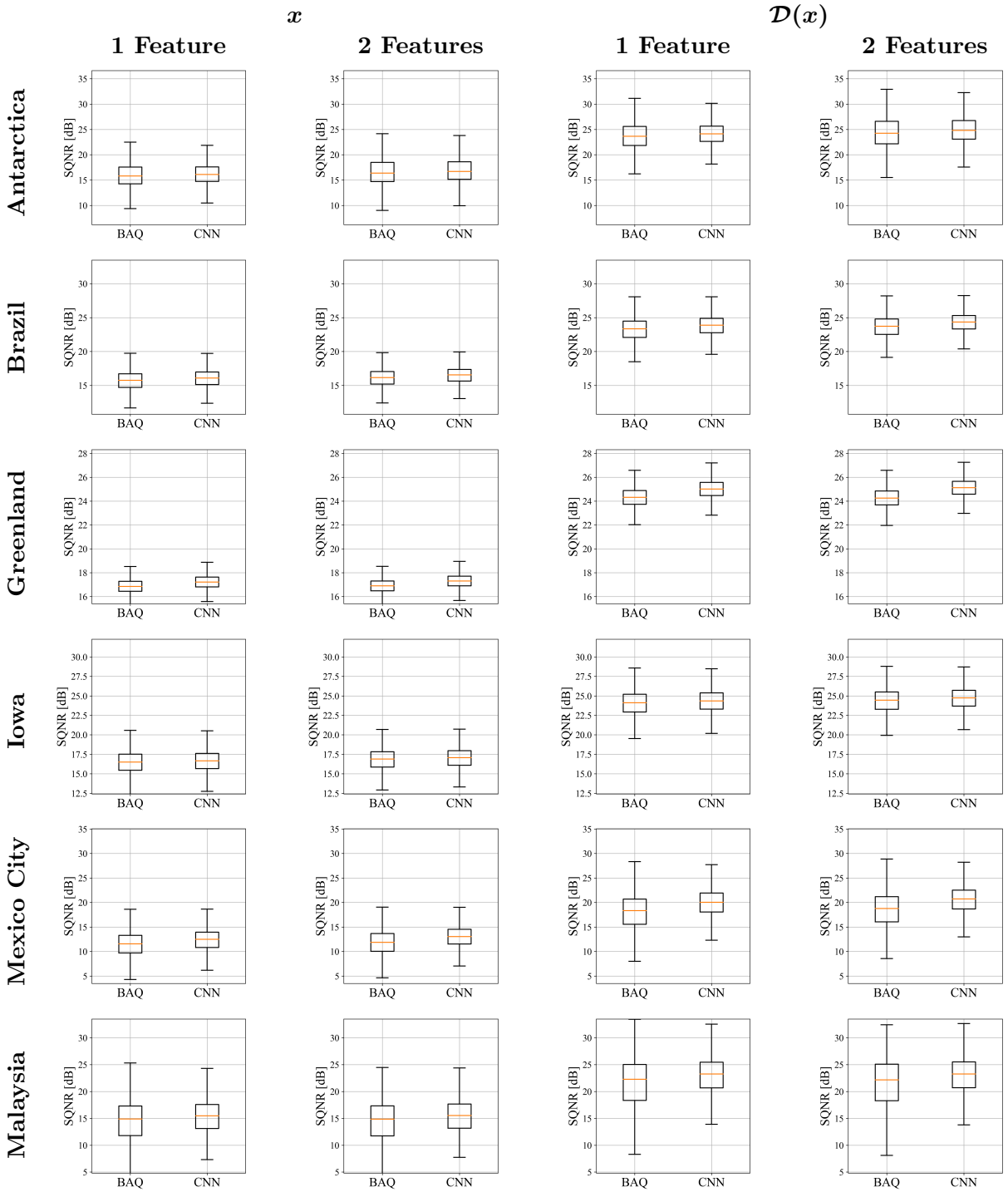



Figure 41: Boxplots of the SQNR computed on the test-set.

the SQNR boxplots reported in Figure 41 and in the SQNR scatterplots in Figure 42.

The scatterplots particularly help to visualize where the biggest performance uplifts take place. The Malaysian on in particular, clearly shows that the biggest gains take place for those samples that start with the lowest SQNR, corroborating previous findings. Finally, the average performance gains reported in Figure 42 confirm that the introduction of the standard deviation feature is beneficial for all tested training-sets, with comparable performance uplifts across the board. In conclusion, the proposed network architecture showed a promising potential for the removal of quantization noise. It

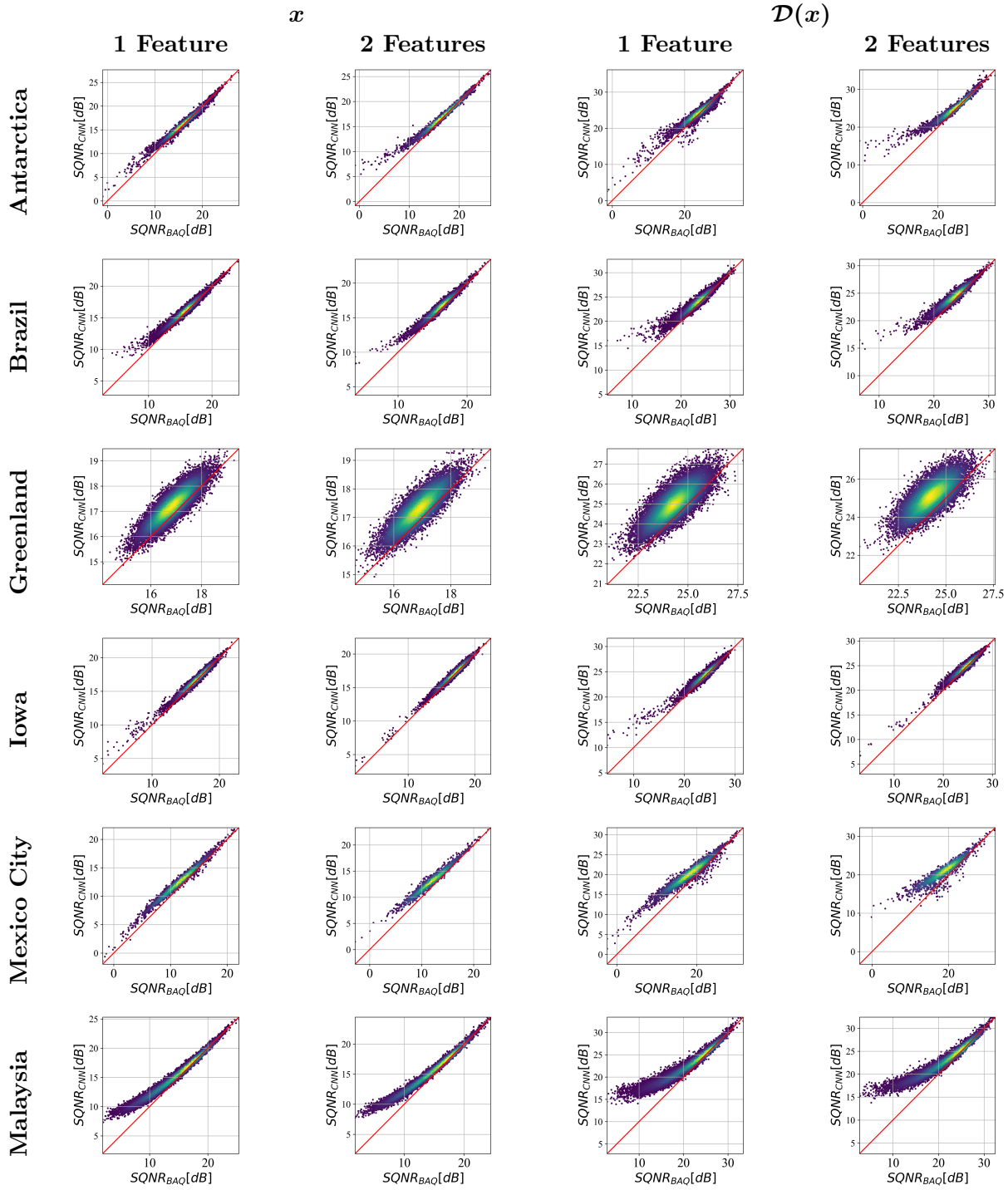



Figure 42: Scatterplots of the SQNR of the input and the denoised samples in the test-set.

appears that the limited number of available acquisition represents a significant limitation for some of the performed experiments, especially for the combined, the Mexico City and Malaysia training-sets. Crucially, the heterogeneous sampling strategy did not cause any significant loss in performance, instead of showing increased gains for that particular test-site, which were limited in their sample populations. It was also observed that the network architecture showed the best uplifts for the patches with the lowest starting performances, with peak SQNR gains of 5 dB. On average, the performance uplifts relies between 0.4 dB to 1.22 dB, when considering the speckle-affected patches, and between 0.53 dB and 2.13 dB for the despeckled ones. The only exception to this is the Iowa dataset, which shows only modest gains between 0.2 dB and 0.35 dB. Overall, the gains are comparable with those obtained by some of the more recent adaptive BAQ algorithms, such as the FDBAQ [2], which is currently utilized within the ESA Sentinel-1 mission. Crucially though, the proposed denoising step is an on-ground computed post-processing step, off-setting the required implementation complexity and computational burden.

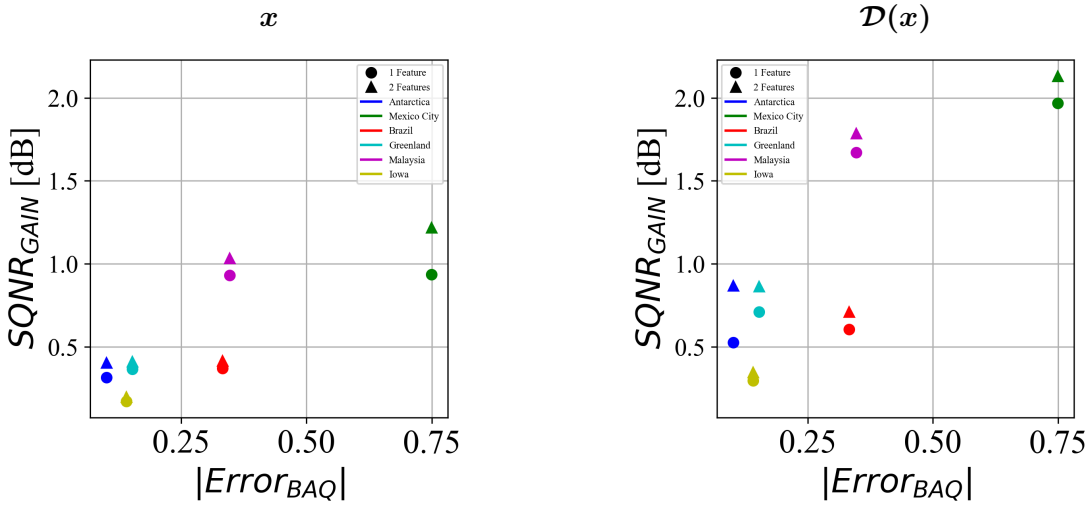



Figure 43: Average performance gain using one and two input features.

# 5 Conclusion and Outlook

In this master thesis, SAR raw data quantization and on-board data volume reduction techniques have been addressed. In the first part of the thesis, some light benchmarks have been performed to asses the different performances of a simple uniform quantizer, an optimum Max-LLoyd quantizer, and the block-adaptive quantization scheme, used by many currently operational satellites, including DLR's TerraSAR-X and TanDEM-X. These showed that, as expected, the uniform quantizer has the worst performance, being sub-optima for Normally distributed signals, like SAR raw data. The optimum quantizer performed the best, at the cost of increased computational demand, while the BAQ compression scheme showed great performance gains with respect to the uniform quantizer, at a fraction of the computational costs of the Max-Lloyd quantizer, justifying in this way its widespread adoption.

Then, the predictability of SAR raw quantization noise has been explored. This work aimed at delivering a simple profiling tool which, starting from a focused SAR acquisition, qualitatively assesses the magnitude of the error introduced by the BAQ algorithm for a particular scene. This entailed the definition of the error correlation window ,associated with the acquisition and focusing processes, and the computation of localized image statistics to estimate the Pearson Correlation Coefficient with respect to a set of performance metrics, namely: the quantization noise magnitude, the signal-to-quantization-noise ratio, the phase difference, and the interferometric coherence loss. This experiment showed that the standard deviation had the greatest linear correlation values with all tested performance and error metrics, concluding that the estimation of the scene-based error metrics is feasible and should be further investigated for even better performance predictions. Based on these results, the proposed data-volume reduction scheme, by means of smart bit-rate allocation, should be further examined as a smart and cheap solution to the ongoing ballooning of on-board storage demand in both existing and planned SAR missions.

Finally, the potentials of a novel but simple raw data end-to-end compression scheme have been explored. The proposed pipeline consists of an aggressive on-board quantization of the acquired SAR raw-data, followed by an on-ground mitigation of the resulting quantization error on the focused data, by applying a deep residual learning-based denoising architecture. The suggested network has been designed to ignore the effects of SAR speckle noise, by integrating a despeckling layer at its output. The resulting architecture has been tested considering combined and individual TanDEM-X acquisitions. The results showed weaker performance uplifts for the combined dataset, but promising gains when individual test-site have been considered. These results suggest that, for the considered test-sites, not enough information has been fed into the neural network to properly model the underlying noise distribution. In order to try to overcome this limitation, the previously-computed standard deviation map has been used as an added input feature to the network, allowing the network to learn additional information on the intrinsic characteristics of the observed area. In this way, better average performance gains of up to 2 dB has been obtained, which are comparable or even better than some of the existing state-of-the-art compression schemes.

Overall, the limited amount of available SAR acquisitions appears to have been the greatest hurdle to successfully train some of the tested networks. Thus, large time-series of multiple test-sites should be considered to drive the development of the architecture forward. Moreover, in order to minimize speckle distortion, multi-objective loss functions should also be investigated in conjunction with more sophisticated despeckling algorithms. As possible further outlook, the performance gains for less aggressive compression-rates should also be explored. Finally, the application domain of the network should be broadened by tackling the problems of phase error and SLC error removal.

# Bibliography


[1] Italian Space Agency. *COSMO-SkyMed Mission and Products Description*, 2019.

[2] E. Attema, C. Cafforio, M. Gottwald, P. Guccione, A. Monti Guarnieri, F. Rocca, and P. Snoeij. Flexible dynamic block adaptive quantization for sentinel-1 sar missions. *IEEE Geoscience and Remote Sensing Letters*, 7(4):766–770, 2010.

[3] J.C. Curlander and R.N. McDonough. *Synthetic Aperture Radar*. John Wiley and Sons, 1991.

[4] J. Fischer, U. Benz, and A. Moreira. *Efficient SAR raw data compression in frequency domain*, volume 4. IEEE, 1999.

[5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[6] Robert M. Gray and David L. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44(6):2325–2383, 1998.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.

[8] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015.

[9] Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 769–776. Curran Associates, Inc., 2009.

[10] Gerhard Krieger, Alberto Moreira, Hauke Fiedler, Irena Hajnsek, Marian Werner, Marwan Younis, and Manfred Zink. Tandem-x: A satellite formation for high-resolution sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11):3317–3341, 2007.

[11] R. Kwok and W. T. K. Johnson. Block adaptive quantization of magellan sar data. *IEEE Transactions on Geoscience and Remote Sensing*, 27(4):375–383, 1989.

[12] Michele Martone. *Onboard Quantization for Interferometric and Multichannel Synthetic Aperture Radar (SAR) Systems*. PhD thesis, Karlsruhe Institut of Technology (KIT), Linder Höhe, 51147, Köln, 2019.

[13] Michele Martone, Benjamin Bräutigam, and Gerhard Krieger. Quantization effects in TanDEM-X data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(2):583–597, 2015.

[14] Joel Max. Quantizing for minimum distortion. *IEEE Transactions on Information Theory*, 6(1):7–12, 1960.

[15] Marvin Minsky and Seymour Papert. *Perceptrons: an introduction to computational geometry*. MIT Press, 1969.

[16] A. Moreira and F. Blaser. *Fusion of block adaptive and vector quantizer for efficient SAR data compression*. IEEE, 1993.

[17] Cristina Párraga Niebla and Gerhard Krieger. Optimization of block-adaptive quantization for SAR raw data. *Space Technology*, 23(2):131–141, 2003.

[18] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11):1338–1351, 2003.

[19] Frank Rosenblatt. The perceptron—a perceiving and recognizing automaton. Technical report, Cornell Aeronautical Laboratory, 1957.

[20] K. Strodl, U. Benz, F. Blaser, T. Eiting, and A. Moreira. *A comparison of several algorithms for on-board SAR raw data reduction*, volume 4. IEEE, 1994.

[21] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman. Learning gaussian conditional random fields for low-level vision. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[22] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.