Data Volume Reduction for Multi-Channel SAR: Opportunities and Challenges

Michele Martone[a], Nicola Gollin[a], Ernesto Imbembo[b], Gerhard Krieger[a], and Paola Rizzoli[a]
[a]Microwaves and Radar Institute, German Aerospace Center (DLR), 82234 Wessling, Germany
[b]Radio Frequency Payloads and Technology Division, European Space Agency (ESA), 2200 Noordwijk, The Netherlands

Abstract

In this paper, we investigate potentials for data volume reduction in multi-channel SAR. These systems allow for high-resolution imaging of a wide swath but, on the other hand, require for their operation the acquisition and downlink of a huge amount of data: together with the intrinsic requirement related to resolution and swath width, this is due to the fact that the effective pulse repetition frequency (PRF) generated by the multiple channels is typically higher than the processed Doppler bandwidth, which introduces a certain oversampling in the azimuth raw data. Therefore, convenient data volume reduction strategies can be proposed, which aim at exploiting the existing correlation between subsequent azimuth samples. In this paper, realistic multi-channel SAR system architectures are considered, and multi-channel raw data are simulated using synthetic as well as real backscatter data from TanDEM-X. We analyse the statistical properties (autocorrelation, Doppler power spectrum, entropy) exhibited by the multi-channel raw signal and discuss the impact of relevant system parameters. Then, possible approaches for data volume reduction for multi-channel SAR are presented, highlighting potentials and limitations.

1 Introduction

Multi-channel SAR allows for overcoming the inherent trade-off of conventional (single-channel) SAR systems in terms of achievable swath width and azimuth resolution. Indeed, by exploiting multiple receiving apertures which are mutually displaced in the along-track dimension, the coherent combination of the individual received signals allows for adequate suppression of the ambiguous parts of the Doppler spectra and, in this way, high-resolution wide-swath imaging is achieved [1], [2]. The drawback for such an improvement of swath coverage and resolution is represented by a significantly larger data volume to be acquired and transmitted to the ground, which poses more stringent requirements on onboard memory and downlink capacity. Therefore, efficient data volume reduction is of utmost importance, as the data rate selected for the digitization of the recorded radar signals directly affects the quality of the resulting SAR products. Currently, one of the most widely used methods for SAR raw data digitization is the Block-Adaptive Quantization (BAQ) [3]. In the last years, starting from the principle of BAQ, novel algorithms have been proposed, allowing for an improved performance and resource optimization, as for the FD-BAQ [4]. These are acquisition-dependent compression schemes, which can be combined with the implementation of non-integer data rates [5]. Furthermore, the potentials of linear predictive coding for efficient SAR raw data compression are shown in [6], [7], and in [8] (the latter in the specific case of staggered SAR systems). In the context of multi-channel SAR, a data-driven data reduction technique based on a principal component decomposition is elaborated in [9], at the cost of a large onboard implementation effort, while the method proposed in [10] aims at exploiting the correlation among the multi-channel azimuth samples by means of a nonadaptive orthogonal transformation (a simple discrete Fourier transform) in combination with an optimized allocation of the quantization rates to the transformed coefficients.

In this paper we investigate and analyse potentials and limitations of data volume reduction for multi-channel SAR. For this purpose, the multi-channel SAR raw signal characteristics are investigated by means of simulation and synthetic data generation, and the trade-space of the relevant SAR system parameters in the context of data volume reduction is discussed. Based on these, possible approaches for efficient data volume reduction in present and future multi-channel SAR systems are discussed.

The paper is organized as follows. Section 2 presents the multi-channel raw data simulations using synthetic data as well as backscatter profiles from real TanDEM-X SAR acquisitions and introduces the metrics (autocorrelation, Doppler power spectrum, entropy) to evaluate the statistical properties exhibited by the SAR raw signal. Section 3 presents the results obtained for three considered multi-channel SAR mission scenarios, and discusses the relevant system parameters which impact the potentials for data volume reduction. Finally, conclusions and outlook are provided in Section 4.

2 Multi-channel SAR Data Analysis

In this section the multi-channel SAR raw data simulations and the investigated parameters are presented.

2.1 Raw Data Simulations

A multi-channel SAR with N receiving azimuth apertures is considered, flying at velocity v_{sat} and transmitting at a pulse repetition frequency PRF_{sys}; dx represents the azimuth separation between the sub-apertures, such that the total azimuth antenna length $L_{\text{az}} = N \cdot dx$. If the following constraint on the transmit PRF is fulfilled

$$\text{PRF}_{\text{sys}} = \frac{2v_{\text{sat}}}{L_{\text{az}}}, \tag{1}$$

then the azimuth raw data stream is uniformly sampled, hence the resulting system is equivalent to a single-channel SAR with $\text{PRF}_{\text{eff}} = N \cdot \text{PRF}_{\text{sys}}$ and in practice conventional SAR processing can be applied. More in general, due to timing constraints, PRF_{sys} is often selected so that (1) is not fulfilled. In this case, an appropriate signal reconstruction of the unambiguous Doppler spectrum of the non-uniform azimuth data, based on the generalized sampling expansion, needs to be carried out on ground by properly combining the N subsampled channels, as in [1], [2]. On the other hand, the Doppler bandwidth of the reconstructed multi-channel signal is larger than the system (transmit) PRF_{sys}, i.e. a finer azimuth resolution is achieved with respect to the corresponding single-channel SAR operating with the full antenna length L_{az} and at the same PRF_{sys}, and multi-channel SAR systems are indeed able to image a swath width of 100 km and more with an azimuth resolution in the order of one meter [2], [11].

In this work, multi-channel raw data have been generated using simulated data as well as backscatter profiles from real SAR acquisitions. In particular, we have considered two simulated azimuth profiles: a homogeneous backscatter profile, where the power level of the complex signal is kept constant, and a profile with a "jump" of 20 dB (i.e. varying from -10 to +10 dB along azimuth, these profiles are not depicted due to space constraints). Furthermore, Figure 1 shows two backscatter profiles obtained from TanDEM-X data: Figure 1(a) depicts an azimuth profile taken from an agricultural (rather homogeneous) site in Iowa (USA), corresponding to the horizontal black line in Figure 1(c); similarly, Figure 1(b) shows a backscatter profile along azimuth over Mexico City, which corresponds to the black horizontal line in Figure 1(d). The area is quite heterogeneous and is characterized by urban settlements and rugged topography. For both profiles in Figure 1, a one-pixel, single-look azimuth line has been considered. The variety of the backscatter profiles and statistics considered in this analysis is exploited to gain more confidence and representativeness with respect to the obtained results and performance. The simulated profiles and the ones shown in Figure 1 are used to generate multi-channel SAR raw data along azimuth, and from these different parameters and quantities are analyzed for three multi-channel SAR system examples, as it is detailed in the next section.

2.2 Parameters for Multi-channel Data Analysis

In this section we present the parameters considered to evaluate data volume reduction for multi-channel SAR: for this purpose the azimuth autocorrelation, the Doppler power spectrum characteristics and the multi-channel signal entropy are investigated, in order to identify the system parameters impacting the potentials for data volume reduction. These parameters are recalled in the following.

2.2.1 Autocorrelation

The autocorrelation $R_x(\tau)$ provides information about the spatial coherence and second-order redundancy of a signal. $R_x(\tau)$ can be expressed as the inverse Fourier transform of the raw data power spectral density $P_u(f)$ as function of the time lag τ. In case of a rectangular antenna aperture, the autocorrelation can be expressed as:

$$R_x(\tau) = \mathcal{F}^{-1}\left\{\sin^4\left(\pi \frac{L}{2v_{\text{sat}}} f\right) \Big/ \left(\pi \frac{L}{2v_{\text{sat}}} f\right)^4\right\}, \tag{2}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier operator, $B_R = \frac{2v_{\text{sat}}}{L}$ is the 4-dB bandwidth of $P_u(f)$. From (2) it can be observed that long antennas L result in more directive beams, implying a more narrow low-pass Doppler filter and therefore higher $R_x(\tau)$ values; also a low satellite velocity v_{sat} implies that target responses are more overlapped, which also results into higher $R_x(\tau)$ values. From the sampled SAR raw data signal $X[n]$, $R_x(k)$ can be estimated along azimuth direction as

$$R_x(k) = \frac{\sum_{n=0}^{N-1} X[n] \cdot X^*[n+k]}{\sum_{n=0}^{N-1} |X[n]|^2}, \tag{3}$$

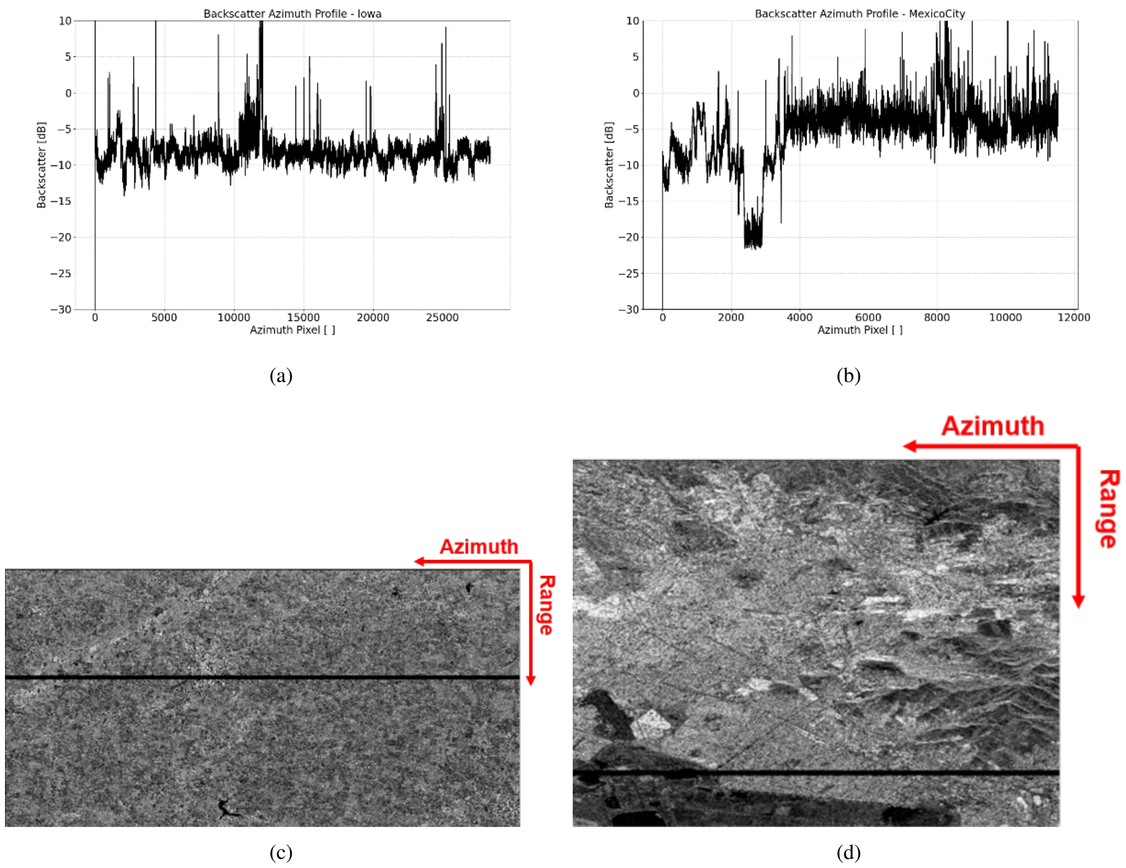
being k the sampled time lag, N the number of signal samples, and * the complex conjugate operator. In this work, $R_x(k)$ is estimated assuming a uniformly sampled multi-channel SAR signal, since a potential correlation/redundancy in the data can be exploited within the N-channel azimuth data block. In general, the azimuth autocorrelation is influenced among the others by the antenna pattern shape (azimuth length, carrier frequency), the number of azimuth channels N, the PRF and, more specifically, the oversampling factor $o_f = \frac{N \cdot PRF_{\text{sys}}}{\text{PBW}}$, being $\text{PBW} = v_{\text{g}}/\Delta\text{az}$ the Doppler processed bandwidth (where v_{g} and Δaz indicate the ground velocity and the azimuth resolution, respectively).

If the multi-channel signal exhibits sufficiently large autocorrelation for small $k > 0$, convenient approaches based on predictive coding and quantization (as proposed in [6], [7], [8] in the context of single-channel SAR) can be exploited for efficient data volume reduction.

2.2.2 Doppler power spectrum

For a multi-channel SAR, the Doppler power spectrum $P_u(f)$ describes the signal power distribution within the N azimuth channels. In particular, the power contribution of the i-th channel ${\sigma^2}_i$ is estimated as the fraction of the corresponding power spectrum integrated over the processed Doppler bandwidth as

$${\sigma^2}_i \propto \mathbf{P}_i \int_{-\text{PBW}/2}^{\text{PBW}/2} \left| \frac{\sin\left(\pi K \frac{f+\text{PRF}(K/2-i)/K}{\text{PRF}}\right)}{\sin\left(\pi \frac{f+\text{PRF}(K/2-i)/K}{\text{PRF}}\right)} \right|^2 df. \tag{4}$$



(a) (b) (c) (d)

Figure 1 Backscatter profiles taken from real TanDEM-X acquisitions used for the multi-channel SAR raw data generation: azimuth backscatter profile taken from (a) an agricultural area in Iowa (USA) and (b) the urban area of Mexico City. The profiles are located in correspondence of the black lines in (c) and (d), respectively.

Hence, ${\sigma^2}_i$ is proportional to the integral of the Dirichlet kernel of the discrete Fourier transform over the processed bandwidth PBW, scaled by the power associated to the i-th transformed coefficient P_i. This, in turn, represents the contribution of the antenna pattern in the corresponding portion of the Doppler spectrum. If the PRF is sufficiently larger than the processed bandwidth PBW (and so o_f is sufficiently large) then the existing redundancy between subsequent azimuth samples can be exploited by applying the method proposed in [10]. Here, a discrete Fourier transform is applied on the multi-channel SAR data block and then a variable bitrate is allocated according to the Doppler power distributions derived in (4).

2.2.3 Multi-channel SAR entropy

A well-known concept of information theory, the entropy measures the average amount of information needed to describe the outcome of a random variable. SAR raw data are modeled as a discrete normally distributed variable X, which can take K values/symbols, and a probability mass $P(x)$, such that its first-order (i.e. single-pixel) entropy can be expressed as

$$H(X) = -\sum_{x=0}^{K-1} P(x)\log_2 P(x). \tag{5}$$

In general, $H(X)$ is maximum if X is uniformly distributed, corresponding to maximum uncertainty and lack of information about the variable's value. Contrarily, the more X is concentrated around a few values (resulting into higher predictability), the smaller $H(X)$, i.e. the lower the information content. To estimate the signal entropy, we consider the ADC-quantized multi-channel SAR data at different bitrates. If entropy values lower than the effective number of bits used for compression are observed, this makes data volume reduction by means of entropy (e.g. Huffman) coding possible.

3 Results and Discussion

In this work, three multi-channel SAR system examples are investigated, and the corresponding system parameters are summarized in Table 1: in particular, different radar frequencies are considered, and the parameters represent realistic scenarios for currently investigated multi-channel SAR missions, such as ROSE-L or HRWS systems at C and X band.

As an example, an HRWS C-band single-platform multi-channel SAR operating in C band (mid column in Table 1) and equipped with $N = 8$ azimuth channels is considered. Figure 2 shows the antenna patterns (Tx in red, Rx

Parameter	ROSE-L	HRWS C-band	HRWS X-band
Carrier frequency, f_c	1.25 GHz (L band)	5.04 GHz (C band)	9.65 GHz (X band)
Satellite height, h_{sat}	697 km	697 km	514 km
Antenna type	Planar array	Planar array	Planar array
Azimuth antenna length, L_{az}	11 m	12.8 m	5.04 m
Number of azimuth channels, N	5	8	[4, 8]
Pulse repetition frequency, PRF_{sys}	1364 (uniform sampling)	1188 Hz (unif.)	3017 Hz (4 ch., unif.) 6035 Hz (8 ch., unif.)
Target azimuth resolution, Δ_{az}	[10, 20] m	[1.5, 3, 5, 10]	[1.5, 3, 5, 10] m
Total processed bandwidth, PBW	2048 Hz	[4691, 2346, 1407, 704] Hz	[4691, 2346, 1407, 704] Hz
ADC Resolution	8 bits	8 bits	8 bits

Table 1 Multi-channel SAR system parameters.

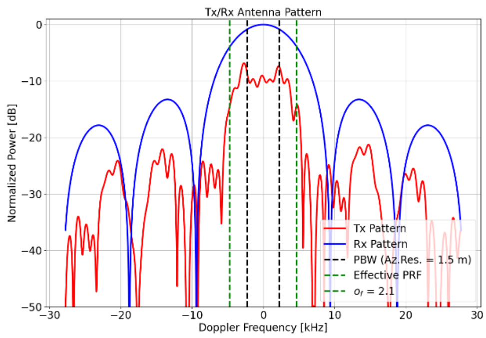


Figure 2 Azimuth antenna patterns (Tx in red, Rx in blue) for the HRWS C-band system. The processed bandwidth, the effective PRF and the resulting oversampling factor (reported in the legend) are depicted for the 1.5-m azimuth resolution case.

in blue) for the considered system. The processed bandwidth and the oversampling factor o_f are depicted for a 1.5-m azimuth fine resolution case. Clearly, in case coarser resolutions are targeted (i.e. smaller Doppler bandwidths), opportunities for increased data volume reduction can be exploited.

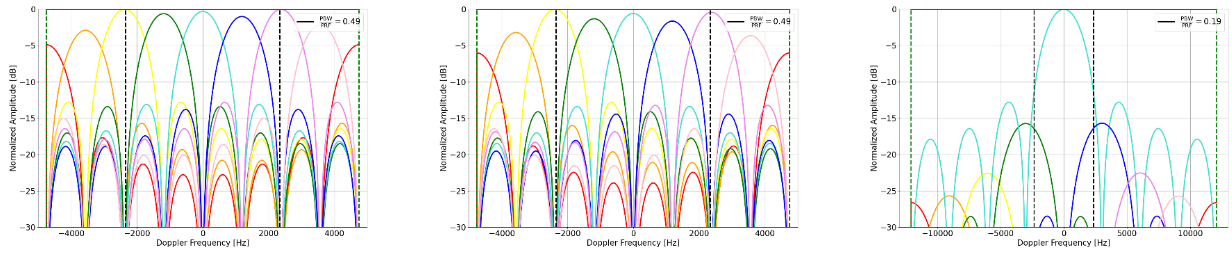
Figure 3 shows the Doppler power spectrum for the C-band multi-channel SAR for the homogeneous backscatter (Figure 3(a)) and the Mexico City profile (Figure 3(b)). For comparison, the power spectrum of the X-band HRWS systems is also depicted in Figure 3(c)), and in all cases Δaz $= 1.5$ is assumed. In general, the power distribution within the Doppler channels is only negligibly affected by the underlying scene (and so the achievable data reduction performance), and the large oversampling factor ($o_f \approx 2$ and $o_f \approx 5$ in the two cases, respectively) allows for convenient data reduction exploiting transform coding in combination with variable bitrate allocation as in [10], especially for lower resolution cases.

In this scenario, convenient methods for data volume reduction aiming at exploiting the intrinsic correlation among neighbouring azimuth and range samples have been introduced. In particular, in [10] the multi-channel SAR data are decomposed by means of a nonadaptive discrete Fourier transform, while in [9] a data-driven transformation of the multi-channel signal based on the Principal Component decomposition is proposed. Furthermore, an optimized allocation of the quantization rates can be applied to the transformed coefficients, which allows for an increase of the resulting performance for a preselected data rate. Finally, in [6], [7] and [12] the applicability of predictive coding in the context of single-channel SAR is demonstrated, but the technique can be also applied for multi-channel SAR among the azimuth direction in presence of sufficient correlation/oversampling.

As metric for performance assessment, we consider the signal-to-quantization noise ratio (SQNR), defined as

$$\mathrm{SQNR} = \frac{\sigma_s^2}{\sigma_q^2}, \quad \text{with} \quad q = s - s_q. \tag{6}$$

In the above equation s and s_q represents the reference (non-quantized) signal and the quantized one, respectively. SQNR represents the figure-of-merit of a quantizer, and can be linked to SAR and InSAR performance measures, such as the total SNR or the interferometric coherence, as discussed in [8].



(a) (b) (c)

Figure 3 Doppler power spectrum for the HRWS C-band obtained for (a) homogeneous backscatter, (b) the Mexico City profile (Figure 1(b)), and (c) the HRWS X-band system with $N = 8$ channels, assuming homogeneous backscatter distribution. For all cases Δaz $= 1.5$ m is assumed.

3.1 Linear Predictive Coding (LPC)

For the application of linear predictive coding on multi-channel SAR data, the coding gain G_M has been derived in closed form based on the autocorrelation values of the corresponding system as in [12] and, from this, it is possible to derive the achievable data volume reduction $DVR_{\%}$ (considering an average SQNR gain of about 5 dB/bit). Table 2 summarizes the achievable G_M and $DVR_{\%}$ for the considered systems, which highlights the great potential of the proposed method at the cost of a modest additional computational burden (for this analysis, a prediction order $M = 4$ is assumed). The large differences in terms of resulting $DVR_{\%}$ (ranging from negligible values for the HRWS C-band up to more than 65% for the 8-channel HRWS X-band system) is due to the varying oversampling factors (in turn related to the number of channels N) as well as to the specific shape of the azimuth antenna pattern.

3.2 Doppler-Based Transform Coding

The results of the SAR data compression based on joint Doppler frequency transform coding of the multi-channel azimuth block and adaptive bit rate allocation as in [10] are reported in Figure 4, in terms of achievable SQNR (in green) with respect to a standard BAQ (in red) for the HRWS X-band system and different assumptions on the number of channels and target azimuth resolution. The resulting $DVR_{\%}$ ranges from about 5% for the case depicted in Figure 4(a) until 25% for the highly oversampled case in Figure 4(d). For the ROSE-L case, given the typical oversampling factor o_f in the order of about 3 one could expect a similar data volume reduction achievable as for the X-band HRWS system with $N = 4$ channels.

4 Conclusions and Outlook

In this paper we investigate potentials for data volume reduction in multi-channel SAR. For such systems, convenient data reduction strategies are proposed, which aim at exploiting the existing correlation between subsequent azimuth samples. We have introduced the relevant parameters used to assess the statistical properties (autocorrelation, Doppler spectrum, entropy) exhibited by the multi-channel raw signal and we have considered three multi-channel SAR system architectures. On the basis of these considerations, we have simulated multi-channel raw data using synthetic data as well as backscatter profiles from real SAR TanDEM-X acquisitions. We have then introduced suitable approaches for data volume reduction for the considered multi-channel SAR based on Linear Predictive Coding as well as Doppler-based transform coding, highlighting potentials and limitations, and assessed the achievable data volume reduction performance. In this paper we concentrated on systems operating in Stripmap mode, while in future studies efficient data volume reduction will be investigated in the context of other acquisition modes, such as ScanSAR, as to be employed on Sentinel-1 Next Generation (NG), together with the inclusion of higher-order statistics (e.g. cumulants) to better exploit the raw data properties and redundancy.

Acknowledgements

This research was funded by the European Space Agency (ESA/ESTEC) project “Data Volume Reduction for Multi-channel SAR” Contract Nr. 4000140269/22/NL/AS/kk.

5 Literature

[1] G. Krieger, N. Gebert, and A. Moreira, “Unambiguous SAR signal reconstruction from nonuniform displaced phase center sampling,” *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 4, pp. 260–264, 2004.

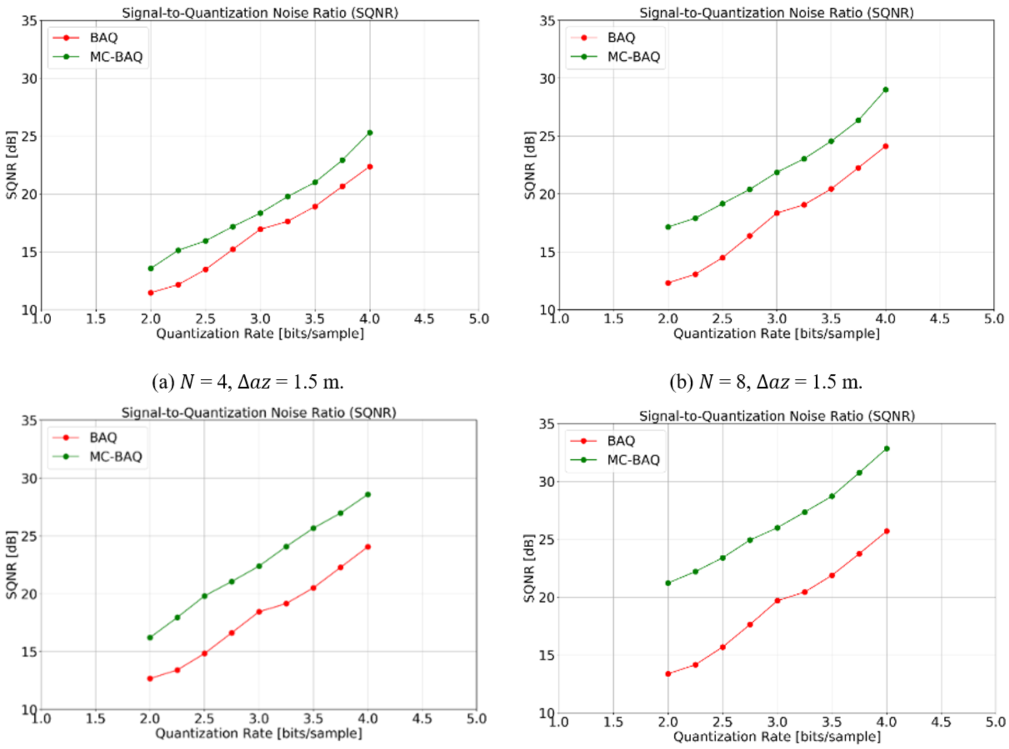
[2] N. Gebert, G. Krieger, and A. Moreira, “Digital beamforming for HRWS-SAR imaging: System design, performance and optimization strategies,” in *2006 IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)*, 2006, pp. 1836–1839.

[3] R. Kwok and W.T.K. Johnson, “Block-adaptive quantization of Magellan SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, no. 4, pp. 375–383, 1989.

[4] E. Attema et al., “Flexible dynamic block adaptive quantization for Sentinel-1 SAR missions,” *IEEE*

System	Frequency Band	N. of channels, N	Prediction Order, M	Coding Gain, G_M [dB]	Data Volume Reduction $DVR_{\%}$
ROSE-L	L	5	4	3.0	14.8
Sentinel-NG	C	8	4	0.4	0.8
HRWS	X	4	4	8.4	41.9
HRWS	X	8	4	13.5	67.4

Table 2 Achievable coding gain G_M and data volume reduction $DVR_{\%}$ using LPC for the considered multi-channel systems.



(a) $N = 4$, $\Delta az = 1.5$ m.

(b) $N = 8$, $\Delta az = 1.5$ m.

(c) $N = 4$, $\Delta az = 3$ m.

(d) $N = 8$, $\Delta az = 3$ m.

Figure 4 SQNR derived for the HRWS X-band system on an homogeneous SAR backscatter quantized with a BAQ (in red), and with the method proposed in [10], as function of the average rate R and for different system assumptions on the number of channels N and azimuth resolution Δaz.

Geosci. Remote Sens. Lett., vol. 7, no. 4, pp. 766–770, October 2010.

[5] M. Martone et al., "Azimuth-switched quantization for SAR systems and performance analysis on TanDEM-X data," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 1, pp. 181–185, 2014.

[6] E. Magli and G. Olmo, "Lossy predictive coding of SAR raw data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 977–987, 2003.

[7] T. Ikuma, M. Naraghi-Pour, and T. Lewis, "Predictive quantization of range-focused SAR raw data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 1340–1348, 2012.

[8] M. Martone et al., "Performance-optimized quantization for SAR and InSAR applications," *IEEE Transactions on Geoscience and Remote Sensing*, 14 Jun. 2022.

[9] Pietro Guccione, Michele Scagliola, and Davide Giudici, "Principal components dynamic block quantization for multichannel SAR," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 2090–2093.

[10] M. Martone et al., "Efficient onboard quantization for multichannel SAR systems," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 12, pp. 1859–1863, 2019.

[11] F. Queiroz de Almeida et al., "A new slow PRI variation scheme for multichannel SAR high-resolution wide-swath imaging," in *2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018, pp. 3655–3658.

[12] M. Martone et al., "Predictive quantization for data volume reduction in staggered SAR systems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, pp. 5575–5587, 2020.